

Shopping behaviour forecasts:
Experiments based on a fuzzy
learning technique in the Spanish
food retailing industry

Mònica Casabayó Bonàs

PhD
The University of Edinburgh
2005

DECLARATION

This thesis is my own work and has been composed entirely by me.

Mònica Casabayó Bonàs

ABSTRACT

The general aim of this thesis is to analyse the possibility of developing synergies when connecting 3 different areas of research namely consumer behaviour, marketing research and artificial intelligence (AI). The three areas of research are very extensive. When analysing the potential links between them, a wide number of triple combinations arise. In addition, the number of combinations can also be increased when applied to different industries but the food retailing industry is selected as the framework of this thesis.

A general overview of the three disciplines is developed. Firstly, consumer behaviour fundamentals are interpreted and reconsidered from a food retailer's perspective. This constitutes one approach to the research in this thesis. Secondly, considering that learning from past data to anticipate shopping behaviours is a retailer's focus of research, an overview of the main marketing research forecasting models and techniques is carried out. Thirdly, machine learning (AI subfield) is also explained in respect of its capability to perform forecasting tasks. Handling customer data is not easy. Information tends to be ambiguous, uncertain and incomplete. Moreover, the customer behaves differently according to his/her situation. Another AI subfield, fuzzy logic (Zadeh, 1965) is also explained as it copes with the concept of partial truth. Having reviewed the three disciplines, the triple combination of 'shoppers (household)', 'forecasting behaviours' and 'fuzzy learning' aspects from each mentioned domain respectively are selected as illustrates the scope of this thesis.

The empirical research consists of two experiments focused on forecasting shopper's behaviour (in particular household shopping behaviour), in the food retailing industry using LAMDA (a fuzzy learning technique). The methodology of research is mainly based on data extracted from a Spanish Food Retailer's (Supermercats Pujol SA) databases. The first experiment is based on LAMDA's supervised learning approach and provides a model to forecast the current customers who are going to defect when a competitor opens a supermarket in the same area. The second experiment is based on LAMDA's unsupervised learning approach and provides a model to forecast the current customers who are going to buy online once the company launches the Website.

Results indicate that marketing expert's judgements are a key point when using learning techniques to forecast behaviours. Customers are not simple robots. People may change their behaviour according to their situation. The results show that when applying the adequacy degree (fuzzy logic concept), the forecasting accuracy increases considerably.

ACKNOWLEDGMENTS

There are several people who helped and encouraged me during these 3 years. Firstly, I would like to extend my warmest thanks to Professors Núria Agell and John Dawson, for their support as doctoral supervisors throughout the project. Their patience, contagious enthusiasm, professionalism and good council have been greatly appreciated.

I would like to extend my thanks to SUPSA, specially to Sr. Pujol and Xavier Roure. SUPSA expressed confidence in me and the research project when part of their customer database was given to me. Further whenever I had some query or doubt, they helped me without question. I am very grateful to Juan Carlos Aguado and German Sánchez for their patience and technical support. My thanks also goes to Professor Jake Ansell for his honest disposition to help and his practical advices.

I would also like to remember Dr. Rob Milne, for his encouragement, curiosity and interest to this project, as in almost everything in life.

Mention I must also go to DURSI (Generalitat de Catalunya) and ESADE for the financial support.

Lastly, but by no means least, and although I know there is no need to thank them again, I really would like to be grateful to my family and friends, not only for their incondicional help, but also always being there. Actually, from now on, I don't think I will be asked the same repeated question 'Com va?' (How is it going on?) again and again. Although I must admit (and I am sure Xavi as well), we will miss that! But this has to do with the near future. By now, I just would like to say:

'Here's looking at you, kid'

TABLE OF CONTENTS

LIST OF TABLES.....	vi
LIST OF FIGURES.....	xi
ACKNOWLEDGEMENTS.....	xiv
CHAPTER 1	1
INTRODUCTION	1
Introduction to thesis	1
1.1 Motivations	2
1.2 Empirical Research	3
1.3 Thesis structure	4
CHAPTER 2	7
CONSUMER BEHAVIOUR FROM A RETAILER'S PERSPECTIVE	7
Introduction	7
2.1 Basics of consumer behaviour.....	8
2.1.1 Definition of consumer behaviour	9
2.1.2 Research orientations on consumer behaviour.....	10
2.2 Managerial interest	15
2.2.1 Why understanding consumer behaviour is important.....	15
2.3 Decision process.....	16
2.4 Determinants affecting decision process	20

2.4.1 Individual factors	21
2.4.2 External factors	21
2.5 Decision Making – Types of decision making.....	23
2.5.1 Merged decision making – Types of customer decision making.....	26
2.6 The importance of identifying the roles	28
2.6.1 The three roles of the customer.....	28
2.7 Customer behaviour fundamentals: A food retailer’s perspective.....	32
2.7.1 Subject of research: The household	34
2.7.2 Research goal	35
2.7.3 Prior data interest	36
2.7.4 Decision making model.....	36
2.7.5 Main key stages.....	40
2.7.6 Standpoint	40
Summary.....	43
 CHAPTER 3	 46
FORECASTING, A MARKETING RESEARCH TASK	46
Introduction	46
3.1 Marketing research. Historical development	47
3.2 Forecasting as a marketing research task.....	49
3.3 The forecasting process	53
3.3.1 Formulating the problem.....	53
3.3.2 Obtaining information.....	53
3.3.3 Selecting and implementing forecasting methods.....	62
3.3.4 Evaluation of forecasting methods.....	65
3.3.5 Advanced forecasting techniques: A general overview	67
3.4 Forecasting customer behaviour.....	71
3.4.1 Behavioural data.....	75
Summary	77

CHAPTER 4	79
ARTIFICIAL INTELLIGENCE SUBFIELDS: MACHINE LEARNING AND FUZZY LOGIC.....	79
Introduction	79
4.1 AI: A general overview	79
4.1.1 Defining AI	82
4.1.2 History of Artificial Intelligence in business	87
4.1.3 A note on the terminology employed.....	91
4.2 Machine learning.....	93
4.2.1 Learning	93
4.2.2 Defining machine learning	96
4.2.3 Machine learning tasks and techniques.....	98
4.3 Fuzzy Logic.....	102
4.3.1 Definition	102
4.3.2 Fuzzy set and fuzzy systems	104
4.3.3 Characteristics of information.....	106
4.4 Fuzzy machine learning hybrid systems	109
4.5 Fuzzy learning applications in marketing	110
4.6 LAMDA, a fuzzy learning algorithm	113
4.6.1 Theoretical aspects of LAMDA	113
Summary.....	117
CHAPTER 5	119
EXPERIMENT 1: FORECASTING CUSTOMER BEHAVIOUR IN THE SPANISH GROCERY INDUSTRY: IDENTIFYING THE CUSTOMERS WHO ARE GOING TO DEFECT.....	119
5.1 General research lines.....	119
5.1.2 Total versus partial defections	122
5.2 Formulating the research problem.....	123

5.2.1 Introduction	123
5.2.2 Setting Research Problem	125
5.2.3 Structuring the problem.....	126
5.3 Obtaining information: Identifying data Sources.....	128
5.3.1 Period of Observation	129
5.3.2 Selecting Data Process	132
5.3.3 Predictors of defection used in this study	135
5.4 Implementation of the forecasting model	143
5.4.1 Data analysis	143
5.4.2 Implementing LAMDA supervised forecasting model.....	147
5.4.3 Validation.....	151
Summary	155
 CHAPTER 6	 157
 EXPERIMENT 2: FORECASTING CUSTOMER BEHAVIOUR IN THE SPANISH GROCERY INDUSTRY: IDENTIFYING THE CUSTOMERS WHO ARE GOING TO BUY ONLINE	 157
6.1 General research lines.....	157
6.2 Formulating the research problem.....	162
6.2.1 Introduction	162
6.2.2 Setting the research problem	165
6.2.3 Structuring the problem.....	165
6.3 Obtaining information: Identifying data sources	166
6.3.1 Period of observation	167
6.3.2 Motivation for online purchasing.....	169
6.3.3 Predictors for ' <i>Shopping Convenience</i> ' used in this study.....	173
6.3.4 Validating corpus	178
6.4 Implementation of the unsupervised forecasting model.....	179
6.4.1 Data Analysis	180
6.4.2 Implementation of unsupervised forecasting model	182
6.4.3 Validation	193
Summary.....	196

CHAPTER 7	198
CONCLUSIONS	198
Introduction	198
7.1 Empirical research's conclusions (Intersection X)	199
7.2 Consumer behaviour and Marketing Research (C1).....	202
7.3 Consumer behaviour and Artificial Intelligence (C2)	208
7.4 Marketing Research and Artificial Intelligence (C3)	209
7.5 Limitations and Further Research	211
 BIBLIOGRAPHY.....	 216

LIST OF TABLES

<i>Table 2.1 Individual factors influencing decision process</i>	<i>21</i>
<i>Table 2.2 External factors influencing decision process</i>	<i>22</i>
<i>Table 2.3 Typology of consumer decision making</i>	<i>25</i>
<i>Table 2.4 Relationship between customer roles and consumer behaviour approaches</i>	<i>30</i>
<i>Table 2.5 Adjustments of consumer behaviour research and customer behaviour research.....</i>	<i>33</i>
<i>Table 2.6 Interpretation of individual personal traits from two different perspectives</i>	<i>41</i>
<i>Table 2.7 Interpretation of environmental factors from two different perspectives... </i>	<i>41</i>
<i>Table 2.8 Interpretation of marketplace changes from two different perspectives ... </i>	<i>42</i>
<i>Table 3.1 Uses of forecasting in some organization function areas.....</i>	<i>51</i>
<i>Table 3.2 Data collection sources.....</i>	<i>54</i>
<i>Table 3.3 Types of forecasting research methods.....</i>	<i>63</i>
<i>Table 3.4 List of possible criteria of evaluating forecasting.....</i>	<i>67</i>
<i>Table 3.5 Table of main contributions of the application of neural networks.....</i>	<i>69</i>

<i>Table 3.6 The most frequently quoted advantages and disadvantages of NN.....</i>	<i>70</i>
<i>Table 4.1 Attributes of an intelligent agent.....</i>	<i>83</i>
<i>Table 4.2 Human attributes related to intelligence</i>	<i>84</i>
<i>Table 4.3 References of NN applications in management</i>	<i>90</i>
<i>Table 4.4 Learning approaches to psychology</i>	<i>94</i>
<i>Table 4.5 Learning and fuzzy learning applications in marketing: Literature review</i>	<i>111</i>
<i>Table 5.1 Average shopping trip frequency.....</i>	<i>131</i>
<i>Table 5.2 Analysis of Saturday shopping trip frequency per week (SSTF).....</i>	<i>131</i>
<i>Table 5.3 Forecasting defections analysis: Literature review.....</i>	<i>134</i>
<i>Table 5.4 Predictors used to predict customer defections</i>	<i>135</i>
<i>Table 5.5 Socio demographic predictor variables selected for the experiment.....</i>	<i>136</i>
<i>Table 5.6 Measuring loyalty and potentiality</i>	<i>138</i>
<i>Table 5.7 Behaviour predictor variables selected for the experiment</i>	<i>139</i>
<i>Table 5.8 Interpurchase variations predictors and the predictive variable</i>	<i>142</i>
<i>Table 5.9 Ruled out predictors.....</i>	<i>144</i>
<i>Table 5.10 Predictors classification according to its type of data: qualitative (non-numerical) and quantitative (numerical)</i>	<i>147</i>

<i>Table 5.11 Example of variations in each customer behaviour (before and after competitor's store opened).....</i>	<i>148</i>
<i>Table 5.12 Quartile distribution from Variable 37 (Variations %).....</i>	<i>149</i>
<i>Table 5.13 Example of the customers' classification according to the purchase variation and quartiles division.....</i>	<i>149</i>
<i>Table 5.14 Interpretation of results based on adequacy results</i>	<i>152</i>
<i>Table 5.15 Interpretation of numerical results based on Maximum adequacy degree</i>	<i>153</i>
<i>Table 5.16 Interpretation of numerical results based on multiple GADS</i>	<i>154</i>
<i>Table 6.1 A summary of convenience attributes from a literature review</i>	<i>170</i>
<i>Table 6.2 Socio demographic predictor variables selected for the experiment.....</i>	<i>174</i>
<i>Table 6.3 Classification of the observed behavioural in-store aspects.</i>	<i>175</i>
<i>Table 6.4 Observed behavioural in-store data predictors selected for the experiment: Selection of the convenience's temporal dimension predictors</i>	<i>175</i>
<i>Table 6.5 Observed behavioural in-store data predictors selected for the experiment: Selection of the convenience's energy/efforts dimension predictors</i>	<i>176</i>
<i>Table 6.6 Observed behavioural in-store data predictors selected for the experiment: Selection of the convenience's spatial dimension predictors.....</i>	<i>177</i>
<i>Table 6.7 Determination of the predictive variable</i>	<i>178</i>
<i>Table 6.8 Variables related to the real online purchasing customer.....</i>	<i>179</i>

<i>Table 6.9 Qualitative and quantitative data</i>	<i>182</i>
<i>Table 6.10 Classifications resulted from LAMDA's unsupervised learning capabilities</i>	<i>185</i>
<i>Table 6.11 Finalist classifications based on unique and stable criteria</i>	<i>186</i>
<i>Table 6.12 Finalist classifications based on the managerial criterion.....</i>	<i>186</i>
<i>Table 6.13 Finalists classifications based on balanced criteria.....</i>	<i>187</i>
<i>Table 6.14 Composition of classification 2, Frank 0.443</i>	<i>187</i>
<i>Table 6.15 Composition of classification 3, Frank 0.454.....</i>	<i>188</i>
<i>Table 6.16 Intervention process applied in classification 2, Frank 0.443.....</i>	<i>189</i>
<i>Table 6.17 Accumulated percentages of the absorbed or lost customers per segment</i>	<i>190</i>
<i>Table 6.18 Classification 2 after the intervention process</i>	<i>193</i>
<i>Table 6.19 Interpretation of results based on a marketing standpoint.....</i>	<i>194</i>
<i>Table 6.20 Measuring LAMDA's forecasts (no overlapping).....</i>	<i>194</i>
<i>Table 6.21 Interpretation of numerical results (no overlapping is considered)</i>	<i>195</i>
<i>Table 6.22 Measuring LAMDA's forecasts (overlapping).....</i>	<i>195</i>
<i>Table 6.23 Interpretation of numerical results (overlapping is considered)</i>	<i>196</i>
<i>Table 7.1 Type of forecasting process carried out.....</i>	<i>203</i>

<i>Table 7.2 Adjustments of consumer behaviour research and customer behaviour research.....</i>	<i>203</i>
--	------------

LIST OF FIGURES

<i>Figure 1.1 Scope of this thesis</i>	<i>5</i>
<i>Figure 2.1 Cognitive and Behavioural learning.....</i>	<i>10</i>
<i>Figure 2.2 Classical conditioning structure</i>	<i>13</i>
<i>Figure 2.3 Reinforcement learning.....</i>	<i>14</i>
<i>Figure 2.4 Model of consumer behaviour.....</i>	<i>18</i>
<i>Figure 2.5 Typology of customer decision making.....</i>	<i>27</i>
<i>Figure 2.6 Roles of customer and research framework.....</i>	<i>30</i>
<i>Figure 2.7 Types of supermarket store choice decision making.....</i>	<i>37</i>
<i>Figure 2.8 Types of product choice decision making when purchasing supermarket basket</i>	<i>38</i>
<i>Figure 2.9 Typology of customer decision making in groceries.....</i>	<i>39</i>
<i>Figure 3.1 Relationship between expected consumer responses and research method designs used</i>	<i>59</i>
<i>Figure 3.2 Schematic conceptions of attitudes</i>	<i>72</i>
<i>Figure 3.3 Types of Behavioural Criteria.....</i>	<i>76</i>

<i>Figure 4.1 IT budgets by segment of the retail sector in USA</i>	89
<i>Figure 4.2 Capital spending on IT by retailers in USA</i>	89
<i>Figure 4.3 Synapse between two neurons</i>	95
<i>Figure 4.4 The neuron</i>	95
<i>Figure 4.5 Representation of crisp sets: TALL or NOT TALL</i>	104
<i>Figure 4.6 Representation of fuzzy set: TALL and NOT TALL</i>	105
<i>Figure 4.7 Examples of numerical and linguistic information</i>	107
<i>Figure 4.8 LAMDA architecture</i>	115
<i>Figure 5.1 Locations of Tremp (Pallars Jussà) and Flix (Ribera d'Ebre)</i>	126
<i>Figure 5.2 Determination of the periods of observation</i>	129
<i>Figure 5.3 Forecasting model based on known behavioural patterns</i>	143
<i>Figure 5.4 Example of quantitative variable:</i>	145
<i>Figure 5.5 Example of the order of magnitude variable</i>	146
<i>Figure 5.6 Interpretation of results based on a marketing standpoint</i>	151
<i>Figure 5.7 Definition of the boundaries for the validation corpus (VC)</i>	152

<i>Figure 6.1 List of channel attributes affecting patronage behaviour</i>	<i>160</i>
<i>Figure 6.2 Evolution of the Spanish B2C e-commerce sales</i>	<i>163</i>
<i>Figure 6.3 Location of Lleida</i>	<i>167</i>
<i>Figure 6.4 Determination of the Periods of Observation (PoO)</i>	<i>168</i>
<i>Figure 6.5 Schematic concept of Shopping Convenience</i>	<i>171</i>
<i>Figure 6.6 LAMDA's unsupervised forecasting model.....</i>	<i>180</i>
<i>Figure 6.7 Results from the application of intervention process in segment 1.....</i>	<i>191</i>
<i>Figure 6.8 Results from the application of intervention process in segment 2.....</i>	<i>192</i>
<i>Figure 6.9 Results from the application of intervention process in segment 4.....</i>	<i>192</i>
<i>Figure 7.1 Scope of the thesis</i>	<i>198</i>
<i>Figure 7.2 Forecasting model.....</i>	<i>200</i>
<i>Figure 7.3 Typology of customer decision making in groceries.....</i>	<i>205</i>
<i>Figure 7.4 Example: Adequacy degree of each customer to the pattern of behaviour 'Stayer'</i>	<i>207</i>
<i>Figure 7.5 Example: Adequacy degree of each customer to the pattern of behaviour 'Total Defector'</i>	<i>207</i>

ABBREVIATION LIST

- AI: Artificial Intelligence
- CBR: Case-Based Reasoning
- CgS: Cognitive Science
- CS: Computer Science
- CSP: Critical Switching Path
- EPS: Extended Problem Solving
- ES: Expert Systems
- GA: Genetic Algorithms
- GAD: Global Adequacy Degree
- GTU: Generalized Theory of Uncertainty
- HDM: Habitual Decision Making
- LAMDA: Learning Algorithm Machine for Data Analysis
- LC: Learning Corpus
- LPS: Limited Problem Solving
- MaxGAD: Maximum Global Adequacy Degree
- MGAD: Multiple adequacy degree
- NN: Neural Networks
- PoO: Period of Observation
- RC: Recognising Corpus
- SL: Supervised Learning
- SUPSA: Supermercats Pujol, S.A
- UL: Unsupervised Learning

CHAPTER 1

Introduction

Introduction to thesis

The aims of this thesis are: 1.To analyse the possibility of developing synergies when connecting 3 different areas of research namely consumer behaviour, marketing research and artificial intelligence (AI); 2. To conduct two experiments focused on forecasting shoppers' behaviour in the Spanish food retailing industry using LAMDA (a fuzzy learning technique).

Consumer behaviour is

'the mental, emotional and physical activities that people engage in when selecting, purchasing, using and disposing at products and services so as to satisfy their needs and desires' (Wilkie, 1994:14).

Marketing research is

'the function that links the consumer, customer, and public to the marketer through information--information used to identify and define marketing opportunities and problems; generate, refine, and evaluate marketing actions; monitor marketing performance; and improve understanding of marketing as a process' (AMA, 2004).

Artificial Intelligence is

'the study of how to make computers to do things which, at the moment, people do better' (Rich and Knight, 1991:13).

Previous research in the three areas is extensive. When analysing the potential links between them, a wide number of triple combinations arise. Moreover, the number of

combinations can also be increased when applied to different industries. Specifically, in the food retailing industry, it is important to note that nowadays, retailers own a large volume of data captured from the scanner systems, loyalty cards programmes and increasing technical advances such as RFID (Radio Frequency Identification) tags, smart cards and Internet servers (Worthington, 1996). Accessibility to this type of secondary data has its pros and cons. Although the majority of retailers, particularly supermarkets, in economical developed countries are able to track their customers' behaviour (Baron and Lock, 1995), they have to face a challenge which consists of finding out the way to turn this raw data into actionable information (Bigus, 1996; Berry and Linoff, 1997) to support decision-making.

Based on that, two experiments are developed in this thesis. These experiments combine forecasting customer's behaviour in the food retailing industry using a fuzzy learning technique in order to deal with the secondary data captured from the loyalty card and scanner systems of a supermarket chain in Spain. The experiments provide evidence of the existing synergies between the three mentioned areas. Experiment 1 aims to forecast customers' defections subsequent to a new competitor store opening. Experiment 2 aims to identify the customers who are presently using a store and are likely to buy online.

1.1 Motivations

The origin of this research started as a result of my personal interest in marketing research. Previous professional experience in marketing and especially in the Spanish food retailing industry stimulated my interest to undertake research on the customer, and his/her behaviour. It was evident that customers are not logical, rational and directed in all cases, and so it is difficult to understand or forecast their behaviour. Based on that, I had some doubts about the effectiveness of some marketing research tools and also their utility to support decision-taking in 'real world' marketing problems.

The food retailing marketplace has reached a mature stage where companies need to be competitive simply to survive (Leeflang and Van Raaij, 1995). Customers are even more demanding and retailers need to design and introduce new ways of learning from them if they are to retain them (Reichheld, 1996a). Accordingly, there has been an increase in the implementation of loyalty cards and similar schemes across European food retailers (Ziliani and Bellini, 2004). The main challenge is not related to capturing data from customers but finding out a means to turn these data into useful, realistic and updated knowledge to support decision making. The most applied marketing research packages are based on statistical methods. However, according to Vellido, Lisboa and Vaughan (1999), there are situations in the marketplace when researchers must deal with uncertain, ambiguous, dynamic and incomplete data. In these cases, applying statistical techniques is not always useful. Related to this, fuzzy logic has the capability to deal with the concept of partial truth and this appears as an interesting possibility to apply to the business field, in particular the retailing industry.

Having noticed and experienced that customers are complex, and that understanding and forecasting customers' behaviour is even more complex, I was curious to find whether there was a more realistic way to understand and forecast customers' behaviour. AI appears to be one possible tool to answer some of my marketing questions raised by the concerns over traditional marketing research methods. Such questions include: 'Is it possible to deal with customer's multiple behaviour? Is it possible to measure Dubois' (1994) chameleonic customer? Is it possible to interpret the marketing research results in a more realistic way? Is it possible to 'humanise' the data to take more realistic decisions?

1.2 Empirical Research

The empirical research consists of two experiments focused on forecasting shopper behaviour (in particular household shopping behaviour) in the food retailing industry using LAMDA (a fuzzy learning technique). The methodology of the research is mainly based on data extracted from a Spanish Food Retailer's internal databases.

The aim of the empirical research is to forecast customer behaviour using just behavioural data stored by the company. Some years ago, the company invested in scanner systems and loyalty card schemes and so owns a large database which captures each customer purchases with the company.

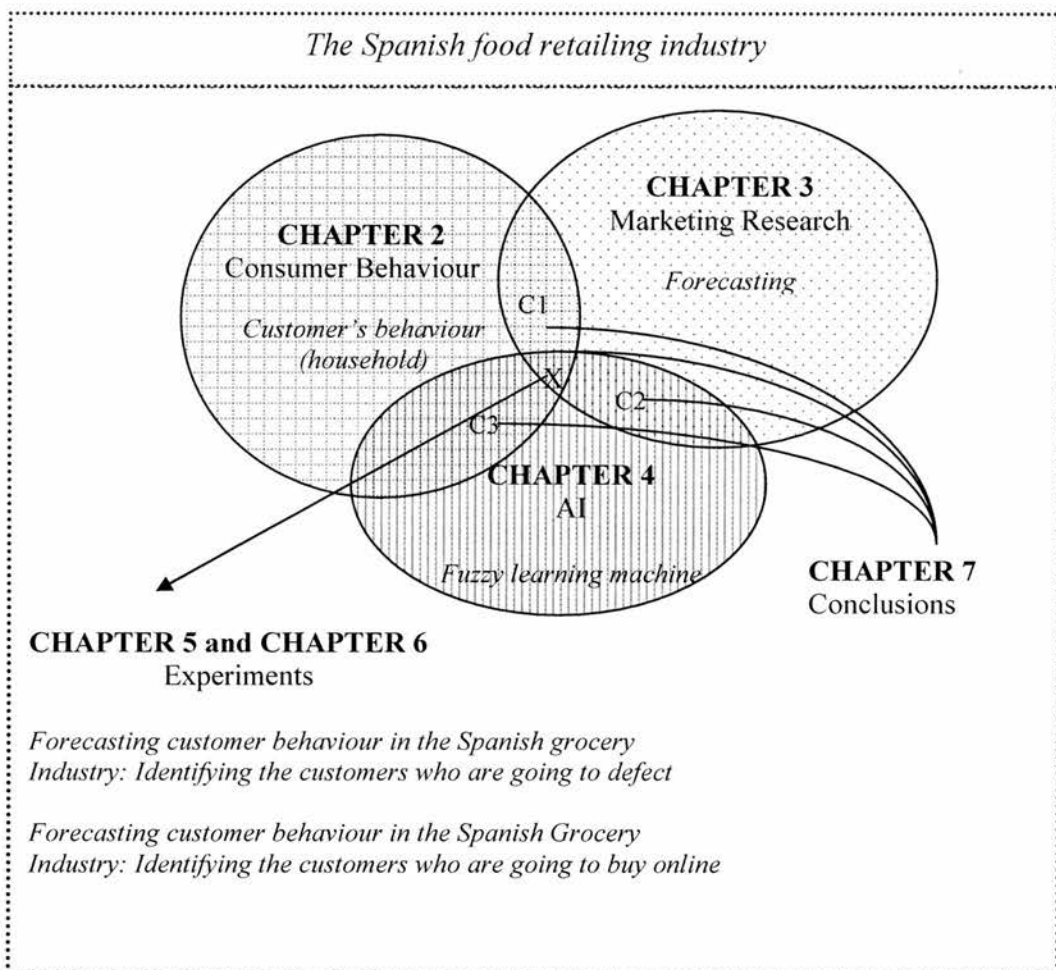
It was evident that the type of data was ambiguous, uncertain and incomplete. Ambiguous because the concepts in marketing are not black or white, and may have different meanings for different persons (e.g. frequency, loyalty, potentiality); uncertain because customer's behaviour is constantly changeable. There are too many variables which may affect the way they purchase. Incomplete because the company had information about the customer, but obviously it was not all the information. Furthermore, LAMDA is a fuzzy learning technique which is able to perform forecasting tasks to deal with fuzziness. In addition, LAMDA has been previously applied in others fields but not marketing.

Based on that, the experiments were developed. The first experiment is based on LAMDA's supervised learning approach and provides a model to forecast the current customers who are going to defect when a competitor opens a supermarket in the same area. The second experiment is based on LAMDA's unsupervised learning approach and provides a model to forecast the current customers who are going to buy online once the company launches the Website.

1.3 Thesis structure

Consumer behaviour, marketing research and artificial intelligence (AI) are the 3 areas of study analysed from a Spanish food retailing industry standpoint. Despite the fact that the three areas of research are extensive, a specific topic of each area is selected to delimit the scope of this thesis. Figure 1.1 illustrates how this thesis has been organized.

Figure 1.1 Scope of this thesis



Chapter 2 reviews the consumer research fundamentals and provides some adjustments when analysing consumer research from a retailer perspective. When analysing the topic from the retailer perspective, the possibility to apply a behaviourist approach arises. Related to that, the terms customer, behaviour, secondary data and purchase decisions precedes the concepts consumer, needs, primary data, preferences and intentions.

Chapter 3 discusses marketing research's goal and provides an overview of the most common methodologies applied in management to forecast customer behaviour.

Chapter 4 provides a general overview of the Artificial Intelligence discipline. In particular, fuzzy logic and machine learning AI subfields are described in detail.

LAMDA, the hybrid technique applied in the empirical research is introduced.

Having reviewed the three disciplines, the triple combination of ‘customers’ behaviour (household behaviour), ‘forecasting’ and ‘fuzzy learning’ aspects from each mentioned domain respectively are selected as illustrated in Figure 1.1. Then, two experiments focused on forecasting shopper’s behaviour in the food retailing industry using LAMDA are presented in Chapter 5 and Chapter 6. The first experiment, Chapter 5, is based on LAMDA’s supervised learning approach and provides a model to forecast the current customers who are going to defect when a competitor opens a supermarket in the same area. The second experiment is described in Chapter 6. Experiment 2 is based on LAMDA’s unsupervised learning approach and provides a model to forecast the current customers who are going to buy online once the company launches the Website.

Chapter 7 summarises and discusses the main findings and contributions of the research.

CHAPTER 2

Consumer Behaviour from a retailer's perspective

Introduction

Most research in consumer behaviour has been analysed from a manufacturer perspective. In the 1960s and until the 1990s, manufacturers had much more power in the marketplace than retailers or other intermediaries (McGoldrick, 2002). Drucker (1992) explains how the real-time information retailers have on purchase behaviour leads retailers to gain power and control in marketing channels.

From a managerial point of view, manufacturers have been more interested in knowing and understanding consumer needs, perceptions and consumption preferences than finding out the shopping likes. It seems that research in consumer behaviour, as an applied field, has been also directed to this line.

Nevertheless, during the 1990s, at the same time that retailers started to have more power in the marketplace, the number of research publications focused on the 'pure' buyer behaviour began to grow (O'Shaughnessy, 1992; Howard, 1994; East, 1997; McGoldrick, 2002; Lempert, 2002; Sheth and Mittal, 2004).

Sometimes knowing the mental and physical activities portrayed by a consumer when using the product are as important as discovering the types of activities when an individual carries out a purchase. Neither the factors which may influence behaviour are the same when he/she buys or consumes. The degree of importance is relative to the researcher's perspective to the consumer behaviour field. For instance, from a manufacturer perspective, the consumer and his/her consumption behaviour are the key focus. Following this manufacturer perspective, the priority of research is to understand the consumer rather than understanding his/her purchasing behaviour. However, from the retailer perspective his/her behaviour is even more important than the individual himself. Based on both manufacturer and retailer perspectives, it was

considered that understanding shopping behaviour precedes an understanding of the individual. There are several reasons which explain this order of preferences.

An overview of the basics and fundamentals of consumer behaviour is provided in this chapter. Then, reconsiderations of these basics are made when developing the customer behaviour fundamentals. It is assumed that there are some particular aspects of the consumer behaviour field that need adjustments when they are analysed from a retailer perspective, in particular food retailer. At the end of the chapter these adjustments are proposed as they are the basis to which it is believed that sometimes, understanding shopping behaviour precedes individual analysis.

2.1 Basics of consumer behaviour

Consumer behaviour is a broad discipline that studies the exchange process involved when individuals or groups acquire, consume and dispose of goods, services, experiences and ideas (Mowen, 1995).

Traditionally, most research in consumer behaviour has distinguished between individuals and companies. However, conventionally, the term consumer has been commonly referred to individuals or household members and customers refers mainly to business units (Sheth and Mittal, 2004). Due to this difference, most of the literature does not make any semantic distinction when referring to the individual who buys, pays or uses the product/service. The term 'consumer' is normally used as a generic term.

Although the study of consumer behaviour is clearly interesting for marketing researchers and managers, (especially since Levitt stated the importance of customer orientation instead of product orientation), it is important to note its attractiveness to other disciplines such as anthropology, sociology, demographics, economics (Devletoglou, 1971), statistics (Armstrong, 2001) and psychology (Watson, 1911; Skinner, 1953; Maslow, 1954; Belk, 1974). Based on that, a close relationship between them is evident across the literature. For instance, theories and models from

sociology, psychology or anthropology are often applied in the consumer research field (Cialdini, 2001). Also consumer researchers are developing their own tools of knowledge which may supplement and retro-feed these social sciences (Foxall, 2001).

2.1.1 Definition of consumer behaviour

Consumer behaviour is

'the mental, emotional and physical activities that people engage in when selecting, purchasing, using and disposing at products and services so as to satisfy their needs and desires' (Wilkie 1994:14).

When splitting the definition into different parts (consumer, sequence of steps and activities), the broad scope of the consumer behaviour discipline appears evident.

Consumer/People

One of the main objectives of consumer behaviour field is to study and understand not only the subject of the action (consumer) but also the features that describe or influence on the individual. Throughout the literature, there is a shared assumption about the necessity to differentiate the general term of 'individual' (Assael, 1987; Mowen, 1995; Kotler, 2000; Sheth and Mittal, 2004). The individual is analysed from two main standpoints. Firstly, the individual is a person or a member of a group, usually a household. On the other hand, the individual is a member of a private business or public organizations. Their general behaviour is so different that sometimes the global term buying unit is found when referring to both of them (Wilkie, 1994). The majority of specialists in this discipline agree that buying may be performed by both a single individual and an organized group.

Decision making process

The process of consumer behaviour implies a set of steps, from acquisition to disposition. Each step has particular features. When investigating the acquisition

phase, researchers focus on analysing the factors that influence the perception and choice. However, the consumption is studied in the following stage of the process. The process is dynamic and changeable. Not only dynamic because both the individual and the environment evolve, but also because it is retro feeding. The process is iterative, recurring and cyclical. Therefore, time must be included when analysing the factors which are likely to appear and influence each stage of the process.

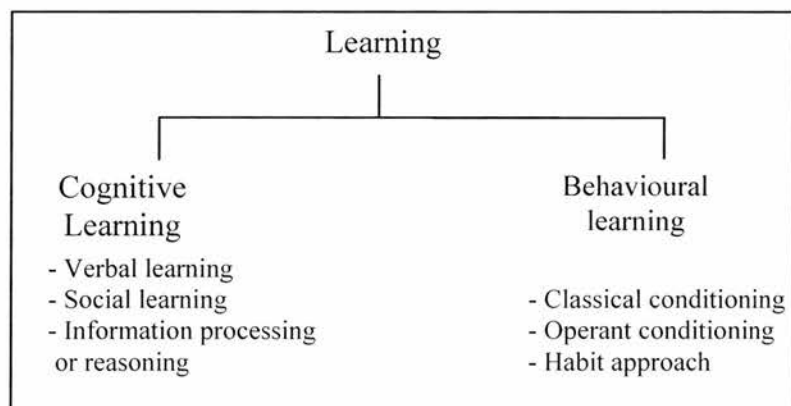
Mental and physical activities engaged (behaviour)

The activities (mental and physical activities) that people initially attempt to undertake in order to obtain a concrete purchasing goal may change. There are specific individual traits, needs and desires that influence the way people act. Moreover, infinite situational circumstances may affect the way a consumer behaves. Although a situation may be defined as a point in time and space, there is the necessity to find the comprehensive taxonomy of situational characteristics that demonstrate their utility when explaining consumer choices (Belk, 1974).

2.1.2 Research orientations on consumer behaviour

Most of the research in consumer behaviour has been oriented towards investigating consumer choice (Mowen, 1995). Consumer choice behaviour has been explained according to two main approaches which coincide with the major distinction in learning theory. These are cognitive and behavioural approaches (See Figure 2.1).

Figure 2.1 Cognitive and Behavioural learning



Despite the fact that the most common approach of consumer researchers and marketing scientists to explain and predict consumer behaviour is extremely cognitive in scope and procedures, when following this perspective, researchers tend to approach problems employing logical empiricist research methods (Mowen, 1995). Both cognitive and behavioural¹ paradigms are explained by Foxall (2001) as follows:

Cognitive approach

This research line is based on the fact that the consumer is a conscious decision maker (Mowen, 1995). Researchers following this cognitive line of research mainly consider the consumer as a receiver and manager of information (information processor) who is moving through rational problem solving which consists of searching for and evaluating the available options which will determine purchase behaviours (Foxall, 2001). Then, purchases are seen as the output of a problem solving process.

As showed in Figure 2.1 the cognitive paradigm of the learning theory includes 3 types of learning. Foxall, Goldsmith and Brown (1998) describe verbal learning as the most related to memory because it is based on verbal repetition. On the other hand, social learning takes place when consumers copy the observed behaviours of others. Finally, information processing is a form of cognitive learning also stressing the role of information. Whatever the cognitive category is, learning occurs when consumers are given specific instructions to follow or deduce which behaviours they should perform based on information.

According to East (1997) the cognitive approach is mainly followed by the majority of popular models in the literature (Engel, Kollat and Blackwell, 1968; Howard and Sheth, 1969). However, despite the popularity of rational and conscious decision making as an explanation for consumer behaviour there are doubts whether it covers much of the field (East, 1997; Foxall, 2001).

¹ Behavioural approach is used by East (1997) to refer to 'behaviourism'.

Behavioural approach

The publications following this behavioural approach are based on the fact that consumers learn from their past behaviour and may use this learning to change or not their future behaviour (East, 1997). Consequently, the behavioural approach explains purchases as the result of learned behaviour.

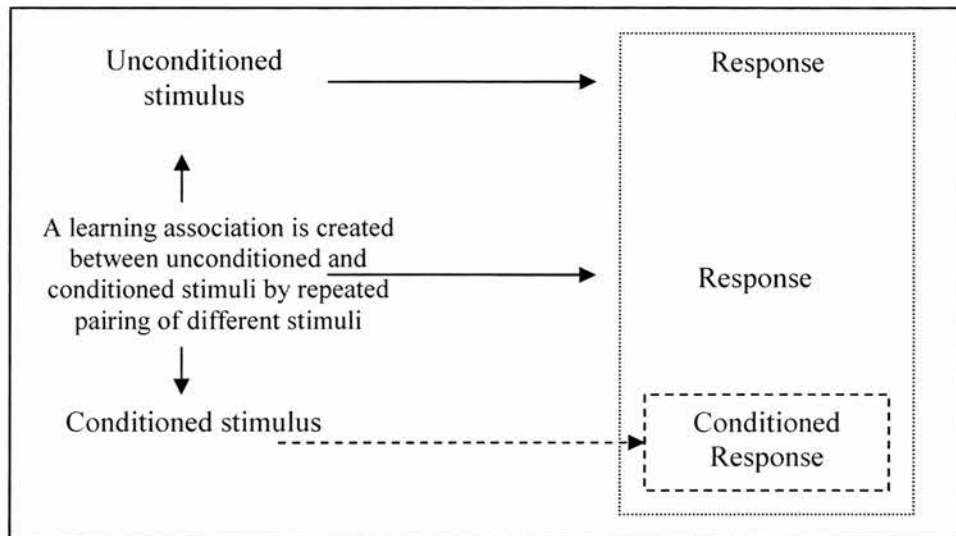
Learned behaviour may be explained according to different standpoints. Classical conditioning and operant conditioning are the most accepted learning theories when giving details of learning. Furthermore, a behaviour setting approach (also called habit approach) is included within the behavioural perspective (East, 1997; Foxall, Goldsmith and Brown, 1998).

Classical conditioning

Classical conditioning was first introduced by the psychologist Pavlov (1927). Classical conditioning has considerable relevance in consumer behaviour, particularly in advertising research, as there are some external stimuli which condition buying particular products (East, 1997).

This behavioural point of view denies the mental process of behaviour to concentrate on how environmental associations manipulate observer responses. As illustrated in Figure 2.2, classical conditioning learning occurs when a stimulus (unconditioned stimulus) that brings out a response is paired with another stimulus that initially does not elicit a response on its own (conditioned stimulus).

Figure 2.2 Classical conditioning structure



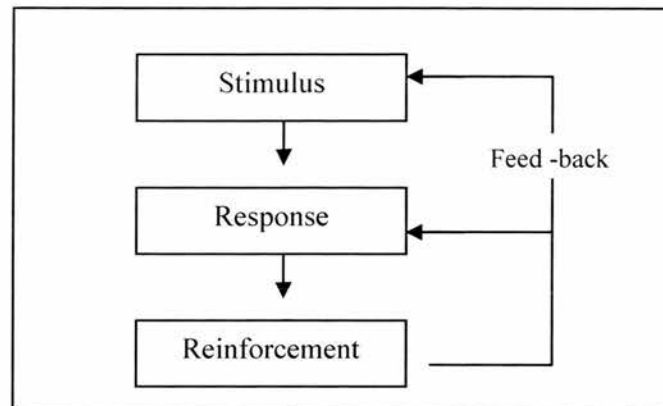
Operant conditioning

Early research in this approach was done by Thorndike (1911) labelled as *'trial and error learning'*. Based on Thorndike's results, Skinner (1953) suggested the term *'reinforcers'* to explain the operant conditioning learning process.

As shown in Figure 2.3, the purchase (the response) may be explained in terms of the experienced rewards and punishments. Reinforcement learning takes place when a consumer learns to behave in a way that boost the positive results and avoid the negative consequences. Based on that,

'the reinforcement approach treats purchase as behaviour learned in response to aspects of the consumer's situation' (East, 1997:7).

Figure 2.3 Reinforcement learning



Source: East, R. *Consumer Behavior: Advances and Applications in Marketing*. Hertfordshire, UK: Prentice Hall, 1997: p15.

Habit approach

Habit appears when an individual behaves in the same way in similar situations. Although problem solving and planning before action are virtually excluded, this does not mean that the person is completely irrational. There are consumers who tend to buy the same brands and to use the same stores over long period of time, precisely because problem solving or planning before action are non-existent. Consequently, a purchase seen from the habit standpoint is seen as

'a pre-established pattern of behaviour, which is elicited by particular situations' (East, 1997:7).

Although the research root and the methodologies used by the cognitive and behavioural approaches are considerably different, some models and theories are shared. According to East (1997), the same theory can explain different facts, or the other way round, one fact can be explained by more than one theory. Based on that, it is important to mention that they are not alternative or opposite approaches but complementary and related. However, at this point, it is important to mention that the behavioural approach is mainly followed within the framework of this thesis.

2.2 Managerial interest

Most of the authors comment on the connection between the foundations of consumer behaviour and marketing strategies (Assael, 1987; Engel, Blackwell and Miniard, 1995; Foxall, 2001). A close link between consumer behaviour research and successful marketing plans is found in the literature.

2.2.1 Why understanding consumer behaviour is important

According to Mowen (1995), consumer behaviour is an exchange process in which different types of resource are transferred between two parties, the consumer and the firm.

Consumer purchases are decided and influenced by individual characteristics, environmental factors, specific situation or circumstances and the product/service he/she needs. There are several types of factors that may affect the way a consumer behaves. Certainly, some of these factors are inherently more difficult to perceive and measure than others. For instance, consumer needs and values are more difficult to establish than demographic features.

Therefore, from a managerial perspective, the multidimensional process by which consumers make purchasing decisions must be understood in order to develop strategic applications. Once the consumer decision process is understood, marketing managers use the information on the consumer in order to define and segment the market, plan marketing strategies, evaluate marketing strategies and assess future customer behaviour (Assael, 1987; Howard, 1994). Thus, the position of the marketer in the marketplace will influence the type of information related to the consumer that the company is interested in. Moreover, the value of information depends on its range of application (East, 1997).

The research framework of this thesis attempts to keep the link between consumer behaviour and managerial interest, particularly in solving food retailing managerial applications.

2.3 Decision process

Knowing literally everything about the consumer would require such a vast volume of information and so many resources that managers should face lots of difficulties to deal with it. For example, there are some individual characteristics that are more likely to affect behaviour than others. In addition, time can change certain types of behaviour. Moreover, concrete interrelationships between variables may vary routine purchases, etc. Consequently, in order to facilitate the understanding of the complex and dynamic consumer behaviour, to help to identify critical points of the stages of the behaviour process, and to support managers in their decision process of establishing the basis for developing marketing strategies, several alternative models have been developed.

The complexity inherent in understanding consumer behaviour has led to the construction of models of the buying process which indicate the stages through which the consumer passes from the time he/she first becomes aware of the need to the time when a product is purchased, evaluated and repurchased again (Foxall, Goldsmith and Brown, 1998).

In the marketing literature, there are a wide range of models that aim at providing an integrative view of consumer behaviour and explaining its influences (Nicosia, 1966; Howard and Sheth, 1969; Bettman, 1979; Assael, 1987; Engel, Blackwell and Miniard, 1995; Mowen, 1995; Mercer, 1998). The major common aspect in all these contemporary models takes into account both the conscious and unconscious level of consumer experience. Not all the models present the same complexity. Some models are more sophisticated than others.

The general, complex and integrated model of Engel, Blackwell and Miniard's (1995) is chosen as a reference for this chapter². Figure 2.4 represents an extended problem solving model of consumer behaviour. Although the model is focused on the internal

² Despite the fact that this model is well-known as Engel, Blackwell and Miniard model, Loudon and Della Vitta (1993) consider it was originally constructed in 1968 by Engel, Kollat and Blackwell.

decision process, each purchase is seen as a result of the combination of four main aspects. Firstly, the power the external stimuli have. Secondly, information processing also may shape purchases. Information processing refers to the way consumers perceive, understand, learn and remember the information is not trivial. According to Harrell (1986:86) perception is

'the process of recognising, selecting, organizing and interpreting stimuli in order to make sense of the world around us'.

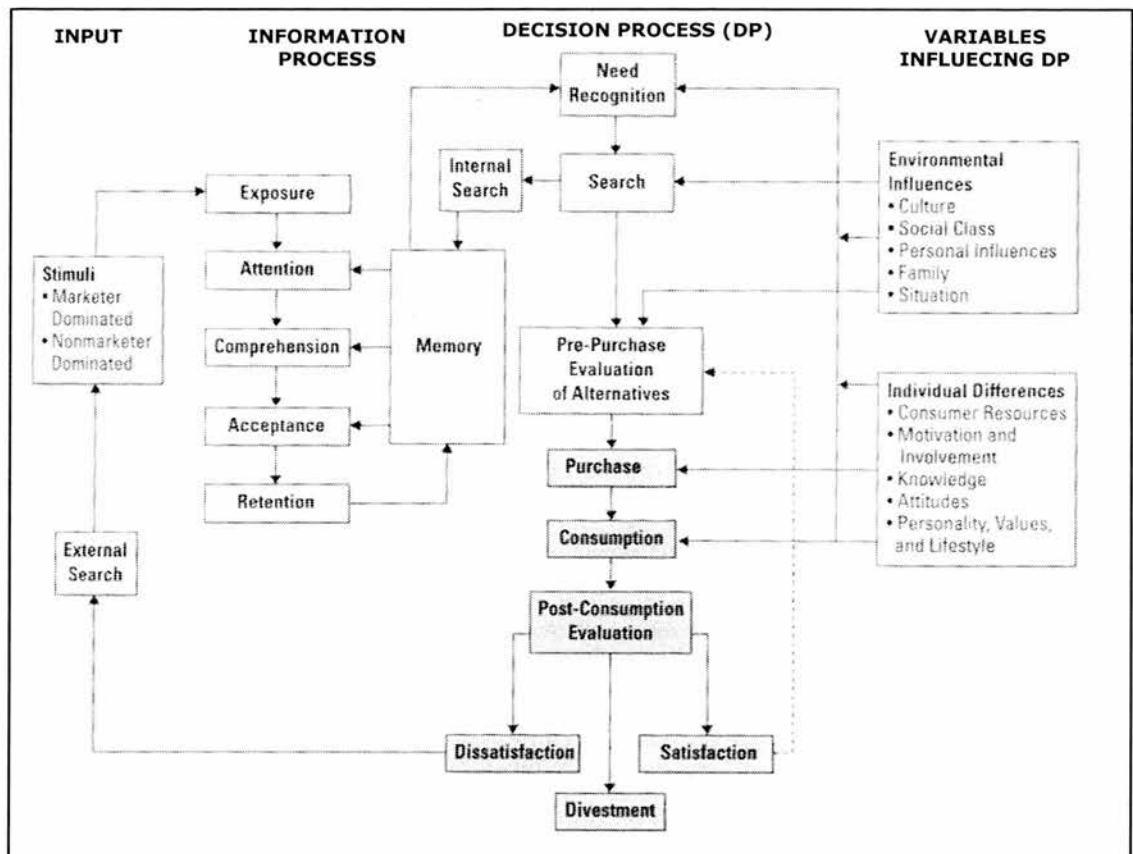
The third aspect is based on the seven stages of the internal decision process. The fourth aspect comprises the variables influencing the decision process. These variables can be split into environmental and individual factors.

As far as the decision process is concerned, consumers pass through seven major stages when making decisions:

- a) need recognition,
- b) search for information,
- c) pre-purchase evaluation,
- d) purchase,
- e) consumption,
- f) post-consumption evaluation,
- g) and divestment.

The authors refer to consumers' behaviour because despite the fact that the model is likely to be applied in store choice (Kotler, 2000), the authors explained the steps performed by an individual when selecting the product to be consumed. Therefore, the explanation of these stages are analysed from a product choice perspective.

Figure 2.4 Model of consumer behaviour



Source: Engel, J.F., Blackwell, R.D. and Miniard, P.W. *Consumer Behavior*. Forth Worth: The Dryden Press, 1995: p154.

a) The recognition of need

The need recognition phase occurs when the consumer perceives a gap between his or her current state and his or her desired state. Once the need has been recognized, the consumer goes about searching and evaluating information to help him or her take a particular decision with reference to a product or service. The need recognition phase in consumer behaviour is where the identification of opportunities occurs on behalf of companies. Identifying a market with unsatisfied desires provides businesses with new sales opportunities.

b) Search for information

Searches for information may be internal or external. Internal searches for

information involve the retrieval from memory or perhaps genetic tendencies whereas external searches include the collection of information from peers, family and the market place. This search can either be passive, where the consumer simply becomes more receptive, or active, where the consumer actively searches in ads, publications and the Internet, amongst others, this is referred to as being market oriented, whereas non market oriented implies sources of information such as friends, family, peers, opinion leaders and the media.

c) Pre-purchase evaluation of alternatives

In the pre-purchase evaluation of alternatives, consumers evaluate the criteria they have on products, services and brands, in order to finally narrow their decision down to resolve to buy one of them.

d) Purchase

This is the phase when consumers become shoppers and decide when, what (product type or brand), where (type of retailer and specific retailer) and how to pay. The act of purchase is carried out.

e) Consumption

It occurs at the point in which the consumer has taken possession of the product. Consumption may be immediate or delayed.

f) Post-consumption evaluation

It occurs when consumers decide whether they are satisfied or not with the product. According to Wilkie (1994:47) satisfaction refers to

'an emotional response to an evaluation of a product or service consumption experience, in which the consumer's expectations are matched by perceived performance'.

This phase has an impact on whether the consumer will undertake a repeat purchase or not. A consumer's satisfaction with the product is determined by five factors:

consumer expectations, actual performance, the comparison between expectation and performance, and the confirmation or disconfirmation of expectations.

g) Divestment

The last stage of the process is divestment. It refers to the process of getting rid of a product or store. This can involve disposal or remarketing, which implies selling the product back into the market.

2.4 Determinants affecting decision process

Understanding consumer behaviour implies understanding the underlying factors that would lead the consumer to decide what, where, when and how he (she) buys (Samli, 1998).

Understanding consumer behaviour is not an easy goal to achieve due to its complexity. According to Zaltman (2003), consumer behaviour is complex because it consists of millions of interactions and dynamic connections between the individual, the context and the product/service which take place in a certain moment of time. Knowing all these infinite multidimensional relationships and controlling some of the dynamic customer variables at the same time is virtually impossible. However, finding out the factors and attempting to measure its influence is a common topic within consumer research field.

Despite the fact that some models are more sophisticated than others, all of them agree with the existence of a range of factors which influence the decision process and help in explaining the reasons why consumers behave the way they do. Usually a major distinction between the individual factors and external factors is found in the literature.

2.4.1 Individual factors

Most authors agree with the view that there is a group of factors directly related to the individual (Howard and Sheth, 1969; Wilkie, 1994; Engel, Blackwell and Miniard, 1995; Mowen, 1995). However, not all the individual features are equally evident. The literature distinguishes between the personal traits and the internal features, which are listed in Table 2.1.

Table 2.1 Individual factors influencing decision process

Individual Factors	
Personal Traits	Internal Features
Gender, age, race, lifestyle, personality	Attitudes, involvement, knowledge, perceptions, emotions, needs and values

The most evident factors are the personal traits such as demographics (gender, age, race), lifestyle and personality. On the other hand, the internal features such as attitudes, involvement, knowledge, perceptions, emotions, needs and values are much difficult to recognise. However, knowing these internal features also help to understand the consumer.

Both personal traits and internal factors are not only necessary to know and understand consumers but also to understand part of their behaviour. The better the consumer is understood the easier is understand and anticipate his behaviour. However, the subject of analysis is important to be defined. As mention later, the subject of analysis in our research is not the consumer as a single person but the household as a customer.

2.4.2 External factors

The decision process is also influenced by external factors.

Table 2.2 External factors influencing decision process

	External Factors	
	Environmental Factors	Marketplace Changes
Static	Social class, reference group, culture	
Dynamic	Role, situation	Demographic, technological, ecological and political changes

According to Engel, Blackwell and Miniard (1995), environmental factors include social class, culture, reference groups (family) and situation (See Table 2.2). The same purchase might be done by an individual or by a group of people. It is important to note that an individual purchase does not follow the same steps as a group purchase. According to Kotler (2000:167),

'a person participates in many groups-family, clubs organizations, and within this participation this person has a role. A role consists of the activities that a person is expected to perform'.

People might play at least five different roles in a buying decision process which are initiator, influencer, decider, buyer and user (consumer). Kotler (2000:176) describes these roles as follows:

*'-Initiator: a person who first suggests the idea of buying the product or service
 -Influencer: A person whose view or advice influences the decision
 -Decider: a person who decides on any component of a buying decision
 -Buyer: the person who makes the act of the purchase
 -User: a person who consumes or uses the product or service.'*

Each role has a direct power to the decision process. Then, finding the role of each member of the group, for instance the role of each family member, may partially explain the final decision.

Social class, culture and type of family are likely to help to understand particular behaviours between different segments but they will not be able to explain multiple behaviours of the same individual. Situations change constantly, the role of the individual may vary considerably but social class, culture and family often remain static. Situations in which consumer behaviour takes place consist of

'all those factors particular to a time and place of observation, which do not follow from a personal and stimulus attributes'
(Belk, 1974).

The same consumer may purchase differently depending on the situation (Belk, 1974; Belk, 1975; Dubois, 1994) and his/her role. Then, the consumer may vary their behaviour just because circumstances changed (Foxall, 1992). Situational determinants are likely to be the most difficult factors to manage because there are a wide range of situations and moments when a consumer may change their normal behaviour. For instance, although a customer usually looks for low prices and bargains, he/she may change his/her mind when guests come at home so the situation led to change his/her buying preferences (Moye and Kincade, 2002).

Moreover, there are other external factors coming from the marketplace which are constantly changing and affect both the customers and the companies. These factors such as demographic changes, technological changes and political changes must be considered as well when understanding and anticipating behaviours (Sheth and Mittal, 2004).

Most of these determinants are dynamic, uncertain and ambiguous. As it is explained in Chapter 4, fuzzy logic allows managing and measuring this type of information and so turning the process of taking decisions into a more realistic framework.

2.5 Decision Making – Types of decision making

Not all the consumers need the same time and same effort for each final purchasing decision. Purchasing decisions do not follow a singular process. According to Engel,

Blackwell and Miniard (1995) the length and depth of the search will depend on several factors including personality, social class, incomes, size of the purchase, past experiences, prior brand perceptions and customer satisfaction. Moreover, according to the personal involvement with the solution (product or service) the difficulty of the purchase, the time and risk invested in the decision process and the number of participants, the type of decision may vary as well. Based on that, a distinction between extended problem solving (EPS), limited problem solving (LPS) and habitual decision making (HDM) for repeated purchased is found over the consumer research literature. Most studies of consumer behaviour use EPS and LPS types of consumer decision making as the prime explanation of consumer choice and do not give enough attention to more habitual aspects of purchase (East, 1997). Others models establish a relationship between the quantities of information used plus the speed of the decision and the stage of the product life cycle (Howard, 1994).

According to Assael's (1987) proposal, there are two dimensions affecting the consumer decision making which are the expansion of the decision process and the degree of involvement of the customer. The former dimension distinguishes between decision making and habit, depending on whether a cognitive and rational search and evaluation process is previously developed or little and virtually not existent information search is carried out. People have habits when they regularly produce much the same behaviour in similar situations (East, 1997). The latter dimension copes with the fact that not all purchases have the same risk or relevancy to each customer. By crossing these 2 dimensions, a 4 cell matrix results as can be seen in Table 2.3.

Table 2.3 Typology of consumer decision making

	High Involvement	Low involvement
Decision making (information search, consideration of brand alternatives)	COMPLEX DECISION MAKING (autos, major appliances)	VARIETY SEEKING (cereals)
Habit (little or no information search, consideration of only one brand)	BRAND LOYALTY (cigarettes, perfume)	INERTIA (canned vegetables, paper towels)

Source: Assael, H. *Consumer Behaviour and Marketing Action*. Belmont, California: Kent Publishing Company, 1987: p13.

Examples illustrated in Table 2.3 are important to be noted as they are relevant to the empirical aspects of this study. All of them, (autos, major appliances, cigarettes, perfume, canned vegetables or paper towels) are product categories. Thus, it is implied that the four types of consumer decision making resulting from the matrix are mainly related to product choice.

Store choice is referred to by Kotler when he describes the types of consumer decision process. According to Kotler (2000), 'complex decision making' goes through 3 stages. Firstly, beliefs about the product/stores are developed. Then, attitudes are built, and finally there is the rational choice. This type of decision appears when consumers have a high involvement with the product and are aware of significant differences. The opposite type of consumer decision is called 'Inertia'. Inertia appears when the stores or products are very similar and when the personal involvement in this category is low. Kotler (2000:177) called it '*habitual buying behaviour*'.

Sometimes, the consumer is very involved with the category of store/product but he does not find it difficult to make a final choice. This type of consumer decision making is called 'brand loyalty'. According to Assael (1987), it appears when the consumer is satisfied with a particular brand and purchases it consistently. Kotler (2000) modifies Assael's matrix considering that this type of decision could be called

‘dissonance-reducing buying decisions’, because these kind of decisions help customers to feel good about the made purchase.

When the category is not very significant to the consumer but he/she perceives lots of differences with the alternatives which make the decision process more complicated, then he deals with the ‘variety seeking’ type of buying decision. Variety seeking is likely to lead to a quick shift of brands because the customer feels he has little to lose. Low involvement can lead to inertia, when customers do not choose a store or product for loyalty but because they do not perceive differences between brands and therefore they still purchasing what they think is a good alternative. Assael (1987: 14) states

‘under low involvement conditions, the consumer will be content to buy a satisfactory brand rather than spend the time and energy searching for the best brand on market’.

The real issue is to find out whether customers are satisfied with the company or, on the contrary, they do not have other alternatives to choose at the moment. Apart from this possible ‘false satisfaction’, a differentiation between categories of products seems to be relevant as well.

2.5.1 Merged decision making – Types of customer decision making

It is important to remark that the typology of decision making found in the literature is focused either on product choice or on store choice but not considering and describing both types of decisions together.

It is believed that every time a consumer takes the decision of purchasing a product (product choice), he/she also takes the decision of where purchasing this product (store choice), and not always the type of decisions of store and product coincide. As it explained in section 2.6, an individual may play different roles. When there is a shopping/ buying tasks, the consumer become a customer. Based on this role play and not only taking Assael’s typology of consumer behaviour into account but also

considering that it can be analysed from store or product decision standpoint (Kotler, 2000), 16 possible scenarios arise (See Figure 2.5):

Figure 2.5 Typology of customer decision making

Product Choice Store Choice	Complex Decision	Brand Loyalty	Variety Seeking	Inertia
Complex Decision	Complex Decision			
Brand Loyalty		Brand Loyalty		
Variety Seeking			Variety Seeking	
Inertia				Inertia

The types of customer decision making can be the same when selecting the product or the store. These cases are located in the diagonal of the table. For instance, a customer may have brand loyalty to a product and a retailer brand as well.

However, there are other situations when the involvement or the time and sources spent when deciding a product or selecting a store are not the same. The green cells show when there is a high involvement with the category of product and a low involvement with the store. For instance, a perfume is likely to be located in the green as generally, the product is likely to be more important than the store. On the other hand, the red cells shows when there is a high involvement with the store but a low involvement when selecting the product. For instance, for a frequent customer, the fact that Tesco runs out of Ariel, does not mean that the customer is not buying any other washing machine gel or that he/she changes to another store just for the product. The yellow cells indicate a high involvement with both the product and the store, although the search of information performed by the consumer is not the same. In general, buying a car, would be located in the yellow cells. Finally, the blue cells show the low involvement of the consumer when deciding the product and the store. For example, the petrol station and type of petrol are located in these cells. However,

the time spent when searching information is different for selecting a store or a category of product.

The purchase act does not consist of a single decision (Foxall, 1992). It is a complex selection involving sub-decisions regarding time and place of purchase, mail or store, types of payment, consumer circumstances, product brand image, past experiences. Based on that, not always the store decision process and the product(s) decision process coincide. There are several variables which may boost one type of decision or another. Also the type of decision when selecting the store may affect the type of decision when purchasing its products and vice versa. For example, the influence of product brand image and consumer attitudes when selecting a store (Collins and Lindley, 2003) are likely to be different from each customer. Moreover, the decision process may vary when choosing between online or of line stores, mostly in groceries. (Degeratu, Rangaswamy and Wu, 2000), mainly due to the difference of involvement perceived. The sector might also predetermine the type of decision making (Assael, 1987). Situations involving grocery purchases are explained in section 2.7.4.

2.6 The importance of identifying the roles

As mentioned before, in the marketing literature, the term ‘consumer behaviour’ has been quite often applied as a generic term, although most of the times its meaning was particularly referred to customer behaviour or shopper behaviour. A distinction between the different terms is provided in the next section.

2.6.1 The three roles of the customer

According to Loudon and Della Vitta (1993:5),

‘the term customer is used to refer to someone who regularly purchases from a particular store or company’.

From Loudon and Della Vitta point of view (1993:5),

'the term consumer more generally refers to anyone engaging in when evaluating, acquiring, using or disposing of foods and services'.

In addition, East (1997: 225) considers that

'The consumer is a shopper when he/she uses and evaluates the retail environment'.

Taking these definitions about consumer, customer and shopper into account and having in mind the different roles that might be performed in a decision process some key conclusions can be drawn and considered in this thesis:

- A consumer (user) is anyone who consumes (uses) the product or service, not necessary the same person who buys.³

'a buyer is the person who participates in the procurement of the product from the marketplace' (Sheth and Mittal, 2004:14).

- A payer is the person who finances the purchase.
- A consumer does not become a customer until (s)he buys at least one time in a specified store.

'a customer is a person or unit who plays a role in the consummation of a transaction with the marketer or an entity' (Sheth and Mittal, 2004:14).

- When a customer is induced to spend money at a store, he/she becomes a shopper and buyer (and/or payer).
- A customer may buy for other consumers.
- A customer is a shopper (buyer) when he/she is in the store looking for

³ Although the person who consumes or uses the product or services is also called end customer, this terminology is not going to be used in this thesis.

product(s).

- When a consumer or a customer uses a supermarket or store, then (s)he becomes a shopper.
- The same person may perform more than one role and be a consumer, customer and shopper at the same time.

Managers and researchers need to know whether a person is deciding on what to consume (consumption behaviour), or whether he/she is deciding on what and how to buy (customer behaviour) or he/she is deciding on the way the shopping is carried out (shopping behaviour). As is shown in Figure 2.6, depending on the specific role or series of roles that the same person is performing, their mental and physical activities are very likely to change, therefore, the research framework is likely to change as well.

Figure 2.6 Roles of customer and research framework

Consumer (User) \neq Payer \neq Buyer (Shopper)	Consumption Behaviour
Consumer (User) = Payer \neq Buyer (Shopper)	Consumption Behaviour and Shopping Behaviour
Consumer (User) \neq Payer = Buyer (Shopper)	Customer behaviour and Shopping Behaviour
Consumer (User) = Payer = Buyer (Shopper)	Consumption, ,Shopping and Customer Behaviour

When relating to the research orientations on consumer behaviour, some assumptions are presumed:

Table 2.4 Relationship between customer roles and consumer behaviour approaches

	Consumer Behaviour Approach
Consumption behaviour	CONGITIVE approach
Customer Behaviour	BEHAVIOURAL approach (Operant learning)
Shopping/Buying Behaviour	BEHAVIOURAL approach (Habit and operant learning)

As it is illustrated in Table 2.4, consumer behaviour seems to be better explained by a cognitive approach because of the general connotation of the context and because

of the traditional emphasis to understand the most internal and emotional features of the consumer, his/her perceptions, preferences and beliefs (Mowen, 1995). When the consumer is in the stage of purchasing they become a customer, and the other paradigms may be more appropriate to explain his/her actions, particularly in groceries.

Customer behaviour is likely to be better understood by the behavioural approach because it is partially based on trial and error learning. As previously mentioned, a customer is a consumer that has bought at least once to the same store. According to East (1997), this approach considers that past experiences are able to describe future buying behaviour. Hence, this approach believes with the fact that managers can vary customer's actions and experiences because they have the power to modify the frequency of actions by increasing the number and the nature of the shopping rewards or reinforcing factors (Skinner, 1953). For instance, retail managers are able to manage the store layout giving stimuli and rewards to the customers. According to Foxall (1992), managers are able to deal with numerous sorts of reinforcements which may partially affect and change customer's situational aspects.

Shopping behaviour is also probably better explained by the behavioural paradigm rather than the cognitive approach mainly because of the exclusion of problem solving. According to East (1997), shopping behaviour may be analysed from a reinforcement learning perspective which explains why promotions have a direct influence to the supermarket basket but also it can be understood from the habit standpoint as helps to explain routine purchases.

From now on, traditional consumer behaviour research found throughout the literature is analysed from a narrower standpoint. The distinction between individual roles and the prime research interest in analysing behaviours make still more evident the suitability of consider behaviourism (behavioural) fundamentals within this thesis.

2.7 Customer behaviour fundamentals: A food retailer's perspective

'Retailing is defined as the set of activities involved in selling products and services to the ultimate consumers (individuals and households)' (Mulhern, 1997: 105).

Over the last two decades there has been a considerable change in retailing (Kotler and Armstrong, 1994; Bareham, 1995). The position of retailing has increasingly changed from mainly playing a passive role with the manufacturers to become the initiator of added value activities in the economy (Dawson, 2005). Notice that the changing role of retailing is particularly manifested in the food industry (Bareham, 1995), which in Europe is in an advanced and mature stage (East, 1997). Consequently, the retail marketing academic research that historically was principally focused on distribution activity has turned into a customer oriented focus of research. Nowadays, the application of advanced information and communication technologies is facilitating economies of scale, growth in firm size and added value to retailers (Dawson, 2005). Thanks to these advanced IT applications a description of the ways technology is revolutionising retailing is explained by Achabal and McIntyre (1987). The initial lack of information that used to face the retailer, has turned into broad access to several types of data. Based on this new context, a customer centred orientation that encompasses a data driven approach is required to effectively manage customer relationships (Mulhern, 1997).

As mentioned before, an overview of the basics of consumer behaviour is needed before reconsidering the possibility to adjust some of the traditional consumer research fundamentals to the retailing field, in particular food retailing. Table 2.5 lists and describes the consumer behaviour basics and proposes some adjustments when developing what we call 'customer behaviour research'.

Table 2.5 Adjustments of consumer behaviour research and customer behaviour research

	Consumer behaviour research (cognitive approach)	Customer Behaviour research (behavioural approach)
Subject of research	Consumer	1. Customer (regular shopper) Household
Research goal	Understanding and anticipating consumer behaviour	2. Understanding and anticipating customer behaviour
Prior Data interest	Consumer insights (primary data)	3. Customers behaviour (secondary data)
Decision making model	Based on product choice	4. Based on store choice and product choice
Main Key stages	Search/pre-purchase evaluation/ consumption	5. Purchase and post-purchasing evaluation
Main standpoint	Manufacturers	6. Retailers

The central point of the consumer behaviour discipline seen from a retailer's eyes is oriented on knowing, understanding and anticipating the way and reasons why people shop, purchase or buy. Then, from the retailers' perspective, the main research focus of analysis is the regular shopper instead of the consumer. Consequently, it does not really matter whether the buyer is also the consumer (user) of the product or not. Neither is it important whether the buyer is the payer or not, understanding and anticipating customers' behaviour is the main interest of research. Based on that, when analysing specific purchases the existence of a merged decision making process between the store and products is assumed.

Within groceries, the seven stages of the decision process described by Engel, Blackwell and Miniard (1995) are virtually reduced to two, which include purchase stage and post purchase stage. It should be noted that buying groceries is characterised as a repeated purchase. Then, the analysis of the purchase act (shopping behaviour) and customer tracking (post purchase evaluations) are critical tasks for the retailers when attempting to enhance the customer-supermarket relationship. Based on that, the prior data interest is oriented in access to behavioural data. Therefore, information related to behaviours is one of the most powerful

weapons to compete within the marketplace. Apart from the behavioural data analysis, information from the marketplace is also required. Obviously, external information may be interpreted from several points of view. Also it is important to note that retailers and manufacturers perceive and interpret changes in the marketplace in different ways.

Customer behaviour research fundamentals are assumed within this thesis and described in depth in the next sections.

2.7.1 Subject of research: The household

Not only are regular shoppers are the subject of the research but also their shopping behaviour. It is important to note that when analysing groceries, irrespective of who has the responsibility for the purchase, the essential point is that many products are purchased on behalf of other members of the buyer's household (Foxall, 2001). Then, the household shopping behavioural pattern becomes more interesting than the individual personal traits.

Household is defined as

'the basic unit of buying and consumption in a society' (Sheth and Mittal, 2004:316).

Despite the fact that the term has traditionally referred to families, it should be noted that household nature and composition is evolving. Currently, some countries, particularly in Europe, are increasing the number of non-family households, which are based on a non-blood link between their members. Household size is also changing. For instance, household size in Europe is gradually decreasing due to fewer births, more single families, and adult people remaining single after divorce. Based on that, a higher number of households but smaller size of families tend to be the future trend of household in Europe (Leeftang and van Raaij, 1995).

The composition of a supermarket basket may be partially shaped by the preferences of all the members of a family (Sheth and Mittal, 2004). Also gender roles may determine the final purchase. For instance, in married households, males' level of shopping responsibility depends on the spouse's occupation as well as the type of goods to be purchased (Dholakia, Pedersen and Hikmet, 1995). Apart from gender roles, other household features may also affect the shopping behaviour and purchased basket. According to Blaylock (1989), race, age, size of the household, time availability and participation of each member have a dramatic influence to household's shopping frequency. Notice that not all the features are equally easy to be recognised and not all the variables have the same degree of influence (Bawa and Ghosh, 1999).

From a food retailer perspective, we believe that focusing on the household features instead of the consumer traits (cardholder traits) when attempting not only to understand purchasing groceries but also forecasting customers behaviours is likely to be more realistic so useful.

2.7.2 Research goal

Once the subject of research is defined, the main goal is to analyse the customer, which is the household. As explained later in Chapter 5, there is a line of research followed by practitioners and academics aimed at retaining and enhancing customers instead of capturing new ones (Dawkins and Reichheld, 1990; Reichheld and Sasser, 1990; Rust and Zahorik, 1993; Storbacka, Strandvik and Grönroos, 1994; Reichheld, 1996a; Mozer et al., 2000; Jones, Mothersbaugh and Beatty, 2000; Harrison, 2003). According to Reichheld (1996b), learning from customers is an early step in the process of retention. Then, it can be inherently induced from the statement that information about customer shopping patterns and about their past behaviours are needed to answer questions such as the following:

- What does/ doesn't the customer buy?
- How much money does the customer spend in our store?
- How much money would the customer be able to spend to our store?

- When does/doesn't the customer normally buy?
- How many trips does the customer make to the store?
- What does the customer expect from the frequent store he/she buys?

From a retailer perspective, the future behaviour is not based on consumer declarations or intentions but on past observed and collected behaviours.

2.7.3 Prior data interest

Within this customer research context, the focus on attempting to identify potential needs and wants that new products and services might satisfy (Sheth and Mittal, 2004), has turned into identifying potential customers' behaviour.

Nowadays, marketers are using their internal database to undertake customer research (Hair, Bush and Ortinau, 2003). In this new research context, collecting primary data such as customers' intentions is not always as relevant as tracking and storing secondary data, for instance customers' actions (McQuarrie, 1988; Ziliani, 2000; Ziliani and Bellini, 2004). Currently, the number of firms in retail and other industries capable of analysing store behavioural data are increasing (for instance banks, airlines and supermarkets). Malhotra, Peterson and Kleiser (1999) state that the application of secondary research is a rising trend shared between practitioners and academics.

2.7.4 Decision making model

As mention in section 2.5.1, customers perform a multiple decision process when selecting between the alternative stores and purchasing particular product(s). Each decision process is likely to be different, not only between store and product (McGoldrick, 2002), but also between the different product categories.

Customers make a decision when the retailer store is chosen and they also make several sub decisions each time a product is added in their baskets. The most applied

combination of decision processes when selecting a supermarket store and when purchasing a supermarket basket is proposed in this section.

As far as the store choice is concerned, firstly note that in the grocery sector the largest difference in the way stores are regarded is between convenience stores and supermarkets (East, 1997). Moreover, despite the fact that people face choices when selecting a store for the first time, this is not the case of our study. It should be noted that the subject of research is the regular shopper (household). Household shopping regularity implies repeated action. Therefore, the store is previously selected by the customer and already learning habits are elicited.

Figure 2.7 Types of supermarket store choice decision making

Product Choice Supermarket Store Choice	Complex Decision (extended decision making and high involvement)	Brand Loyalty (habit and high involvement)	Variety Seeking (extended decision making and low involvement)	Inertia (habit and low involvement)
Complex Decision (extended decision making and high involvement)	Complex Decision			
Brand Loyalty (habit and high involvement)		Brand Loyalty		
Variety Seeking (extended decision making and low involvement)			Variety Seeking	
Inertia (habit and low involvement)				Inertia

Then, although complex decision may appear (for instance when selecting a supermarket brand for the first time) and also variety seeking may become visible, (when there is a wide evaluation of store brand due to a unsatisfactory experience, for example), two most common types of customer decision corresponding to store choice are mainly assumed, which are brand loyalty and inertia (See Figure 2.7).

In reference to the product choice (See Figure 2.8), obviously, not all the product categories from a supermarket are equally perceived. For instance, meat purchase has higher involvement than purchasing paper towels. Also, all customers do not share the same degree of involvement in each category. For example, a recent mother

perceives a higher risk of failure when purchasing a specific brand of napkins that a single woman. However, in general terms, the supermarket basket is commonly perceived as a 'low involvement' purchase decision (East, 1997; Buckinx and Van den Poel, 2005). Based on that, variety seeking and inertia are the most relevant types of customer decision process. Based on that, it is important to mention that most non-durable products comprise several brands which are so similar to each other in terms of their basic attributes that customers do not discriminate among them (Foxall, 2001). Usually, customers select from a set of tried and tested brands 'evoked set' or 'choice set'. This means that brand evaluation is accomplished through 'trial and error' and brand loyalty is rarely allocated to a single brand.

Furthermore, there are some products which reappear every time the supermarket basket is filled. Some purchases are made again and again. If we look at the items purchased on the supermarket we see that we have bought the same product and same brand many times before. Based on that, brand loyalty is also possible when product choice when frequent items that require 'little conscious attention' are purchased.

Figure 2.8 Types of product choice decision making when purchasing supermarket basket

<div>Product Choice</div> <div>Store Choice</div>	Complex Decision (extended decision making and high involvement)	Brand Loyalty (habit and high involvement)	Variety Seeking (extended decision making and low involvement)	Inertia (habit and low involvement)
Complex Decision (extended decision making and high involvement)	Complex Decision			
Brand Loyalty (habit and high involvement)		Brand Loyalty		
Variety Seeking (extended decision making and low involvement)			Variety Seeking	
Inertia (habit and low involvement)				Inertia

When matching the most common types of decisions that appear when select supermarket store (Figure 2.7) and product choice (Figure 2.8), the following 6 scenarios arise.

Figure 2.9 Typology of customer decision making in groceries

Product Choice Store Choice	Complex Decision (extended decision making and high involvement)	Brand Loyalty (habit and high involvement)	Variety Seeking (extended decision making and low involvement)	Inertia (habit and low involvement)
Complex Decision (extended decision making and high involvement)	Complex Decision			
Brand Loyalty (habit and high involvement)		Brand Loyalty		
Variety Seeking (extended decision making and low involvement)			Variety Seeking	
Inertia (habit and low involvement)				Inertia

The regular shopper has a preferred store or set of preferred stores. A number of studies show that households have a routine of supermarket shopping which often includes one weekly main trip and one or more secondary quick trips (Kahn and Schmittlein, 1989) to the same store or a evoke set of stores (Foxall, Goldsmith and Brown, 1998). Shopping trips to the selected store may have been motivated by a relative loyalty to the brand or just for inertia. Whatever is the involvement or motivation, the type of decision when purchasing the basket of products may be based on brand loyalty, variety seeking or inertia as well. Consequently, despite the fact that a brand loyalty may be found in the store choice and in all the product categories, and also inertia can appear in both choices, there are other possibilities which can appear as well. For instance, a regular shopper may have a strong relationship with the supermarket store and a low involvement with some product categories. Also he/she can have a low involvement with the supermarket brand and also shows a strong preference for a specific product brand.

As illustrated in Figure 2.9, shadowed cells show the typology of customer decision making when groceries are purchased by a regular shopper. As it is shown, 6 out the

initial 16 scenarios resulted from merging store choice and product choice decisions are considered.

2.7.5 Main key stages

An aim of this thesis is to analyse shopping behaviour once the customer has already decided to buy regularly in a particular food retailer store. Therefore, within this context of research based on learning behaviour, it is clear that not all the stages from the cognitive decision making process previously introduced in section 2.3 are equally relevant. The key stages of the mention consumer model provided by Engel, Blackwell and Miniard (1995) have been reconsidered and limited to purchase and post purchase evaluation.

2.7.6 Standpoint

As postulated by the cognitive approach, understanding customer behaviour is partially achieved when knowing the customer individual variables. The personal traits and internal features found in the consumer behaviour literature (explained in section 2.4.1) may be useful within this customer behaviour research context. However, we would like to highlight that the same information may be interpreted differently according to the manufacturer or retailer standpoint. For instance, when analysing the personal traits such as gender, age and race the same data has different interpretations. Table 2.6 illustrates this double interpretation.

Table 2.6 Interpretation of individual personal traits from two different perspectives

Personal Traits	Manufacturer's Perspective	Retailer's Perspective
Age	<ul style="list-style-type: none"> - Some products are delimited to a certain age - Social and emotional values are more important for teenagers - The greying trend in Europe is likely to affect the design of certain products (Leeflang and Van Raaj, 1995) 	<ul style="list-style-type: none"> - Elderly buyers need more service and convenience - Older buyers prefer to buy based on relationship with the seller - Teenagers tend to look for the highest shopping experiences
Gender	<ul style="list-style-type: none"> - Many products are gender specific 	<ul style="list-style-type: none"> - Women in the workforce might seek convenience and time saving in shopping - Men are still learning shopping skills in some cultures
Race	<ul style="list-style-type: none"> - Consumers seek products compatible with their skin - Ethnic tastes in food and clothes differ 	<ul style="list-style-type: none"> - Some ethnic groups seek or avoid some products and suppliers in a store - Race influence the interaction between the store

Source: Extracted from Sheth, J.N and Mittal, B. *Customer Behavior: A managerial perspective*. Ohio, USA: Thomson, 2004.

As mentioned in section 2.4.2, culture, social class or personal situations are some of the variables which may also determine the decision process. Interpreting these variables will help to explain individual changes.

Table 2.7 Interpretation of environmental factors from two different perspectives

External individual Factors	Manufacturer's Perspective (consumer role)	Retailer's Perspective (buyer role)
Culture	<ul style="list-style-type: none"> - Corporate culture constraint what people wear, eat or drive. 	<ul style="list-style-type: none"> - Culture assign the buyer role according to their sex-norms
Social Class	<ul style="list-style-type: none"> - Some products are used among a particular social class 	<ul style="list-style-type: none"> - There are times of shopping shared by a specific social class
Situation	<ul style="list-style-type: none"> - Products are consumed according to the consumer situation (needs) 	<ul style="list-style-type: none"> - In store atmospherics may influence customer's predisposition of shopping (Bacon, 1984; East, 1997; Moye and Kinkade, 2002)

Source: Extracted from Sheth, J.N and Mittal, B. *Customer Behavior: A managerial perspective*. Ohio, USA: Thomson, 2004.

For example, when analysing situational factors from a retailer standpoint, atmospherics become essential. Previous studies demonstrate the powerful influence of atmospherics. This influence is palpable both in off line shopping (Kotler, 1973) and on line shopping (Eroglu, Machleit and Davis, 2001). Aspects such as music, colour, odours, and temperature can vary personal moods, and subsequently, shopping behaviour (East, 1997). Easy surfing, friendly interface and connection speed are likely to affect in shopping behaviour as well (de Kare-Silver, 1998).

In addition, changes in the marketplace are not always perceived in the same way. There are social, ecological and technological changes that directly affect the way people consume or shop. Once more, information provided by these environmental changes is likely to be differently understood, depending on whether it is analysed from a retailer or a manufacturer point of view as it is shown in Table 2.8.

Table 2.8 Interpretation of marketplace changes from two different perspectives

Marketplace changes	Manufacturer's Perspective	Retailer's Perspective
Social	- Save time when consuming (Lavin, 1993)	- Save time when shopping (increasing convenience)
Ecological	- Organic and healthy products - Stationary consumption of products	- Modifying category management - Increasing specific retailing services in specific times of the year
Technological	- Product Customisation	- Virtual shopping

Source: Extracted from Sheth, J.N and Mittal, B. *Customer Behavior: A managerial perspective*. Ohio, USA: Thomson, 2004.

As far as social changes are concerned, many consumers and customers suffer from 'time poverty'. Therefore, 'spare time' is more precious and consumers are searching for products or services that help them to save time (de Kare-Silver, 1998; Bowlby, 2000). Moreover, Griffith and Krampf (1997) point out that the culture of convenience is already being evidenced in the retailing sector by decreases in times spent in store visits. Also Fram and Axelrod (1990) share the change in shopping

habits due to the lack of time. The authors reported that customers want to reduce the amount of time spent shopping because some customers felt that shopping had added stress to their lives (Aylott and Mitchell, 1998).

In reference to ecological changes, environmental awareness and the climate are highlighted. Over the past two decades, growing environmental awareness and concerns about health and quality of diet have led many people to question modern agricultural practices. Despite the fact that consumer awareness of organic food is increasing, not all the supermarkets are responding, and there is still a gap between supply and demand, in particular in UK and Germany (Latacz-Lohmann and Foster, 1997). The climate also can influence consumer or buyer behaviour in different ways. Specific types of food are only consumed during a season of the year. Weather conditions may postpone purchases or can lead to asking for more retailing services.

From a technological perspective, de Kare-Silver (1998) described changes in technology directly affect both consumer (increasing the customization of products and services) and customer behaviour (the virtual shopping experience).

Summary

The purpose of this chapter has been to review concepts of consumer behaviour from a retailer perspective, in particular food retailer perspective. Despite the fact that traditionally, much consumer research literature has been developed from a cognitive approach, the appropriateness of a behaviourism approach when analysing customer shopping behaviour in a supermarket is suggested and assumed within the chapter.

Moreover, it is shown that consumer decision making is not a singular process. Different types of consumer decision making are found over the literature. Depending on the time and energies spent when purchasing a product or selecting a store, the type of decision making may vary. However, both the store choice and product choice are analysed separately. We believe that a final purchase is a result of the combination between a type of decision making when choosing the store, and a type of decision making when choosing the product (s), and vice versa as well. Then,

the necessity to find a range of integrated types of decision making is suggested and a double choice matrix provided.

Particularly in groceries, purchases are commonly characterised to be repetitive or eventually purchases based on habit, as in shopping trips as in the supermarket basket purchased. For example, the recognition of lack of milk, virtually leads to a direct purchase, simplifying the information search and evaluating brand alternatives. Implicit in repetitive purchasing is the assumption that customers learn from past experience and buys the product or/and at the store what is most expectations matching. When this happens, information search and alternatives evaluation are limited or non existent since the customer has decided to buy almost the same basket again. Therefore, the decision making is considerably simplified. The lack of complex decision making is important to be remarked as it shows that when studying groceries, customer past experiences and reinforcement learning become decisive concepts. Hence, the habit approach is also considered.

Furthermore, the consumer as central to study has been replaced by the customer and his/her roles. A distinction between the roles of the customer is highlighted, and the interest in the regular shopper is proposed as the subject of the research.

The traditional research goal on understanding the individual to describe and anticipate his/her behaviour is replaced by the aim of understanding and anticipating the household's shopping behaviour. Knowing household features becomes more important than recognizing individual features. It is also believed that when understanding regular shopper behaviour, the past behaviour stored in the internal retailer databases is more interesting than the specific individual features of the cardholder. Based on that, we assume the regular shopper behaviour stored over time represents the household shopping behaviour.

Behavioural analysis of consumer choice is under-studied among marketing researchers but it promises to contribute strongly in marketing literature (Foxall, 2001). In addition, researchers using a behavioural measures of habitual choice

generally rely on consumer data because such data permit the researcher to trace the consumers' pattern of brand purchases over time (Assael, 1987). Based on this, it is believed that learning from customer's past shopping behaviour from a Spanish Supermarket chain is the basis to anticipate future shopping behaviours. It is also assumed then that retailers not only have the possibility to access to secondary behavioural data, but also to other types of information available in the marketplace. This information can be interpreted differently, according to a manufacturer or a retailer perspective. A retailer perspective is taken in this thesis.

CHAPTER 3

Forecasting, a marketing research task

Introduction

'Marketing research is the function that links the consumer, customer, and public to the marketer through information--information used to identify and define marketing opportunities and problems; generate, refine, and evaluate marketing actions; monitor marketing performance; and improve understanding of marketing as a process' (AMA, 2004).

According to the definition, marketing research is the process which turns data into actionable information. As a process, several stages exist, from selecting the data collection method to interpreting the results. Firms, and specifically retailers, may choose between several procedures to collect, analyze or interpret data. This choice comes from advances in the technologies employed. Consequently, depending on the marketing research problem to solve, the data, methods, time and resources required are likely to be different.

It is important to highlight that the forecasting framework studied in this chapter is based on a retailing perspective. The research interest in this part of the thesis is to overview the general steps which intervene in a forecasting process and to review the most common applied forecasting methods, from a retailing point of view. According to this research delimitation, marketing research as a whole is not going to be considered. Neither is forecasting research studied in general.

It is important to make a distinction between two main types of forecasting applications in retailing. There is a considerable difference when forecasting sales or when forecasting the store brand behaviour, for instance. Based on this example, a classification between pure statistical numeric forecasting (e.g. sales, prices) and behavioural forecasting (commonly applied when predicting store location, store choice, store layout, cross-selling techniques, targeting actions) is evident. Accordingly, this chapter is particularly focused on behavioural forecasting, as

predicting customer behaviour, is one of this dissertation's fundamental aims. Based on that, a view of the historical development in marketing research is first provided. Afterwards, the stages of the forecasting process are reviewed. Then, short term major forecasts proposals in the food retailing industry are listed. Finally, the forecasting problem related to customer behaviour is introduced.

3.1 Marketing research. Historical development

According to Kinnear and Taylor (1996) the development of marketing research is parallel to the historical development of research methodology in the social science.

Marketing research made major advances from 1910 to 1920. Questionnaire studies or surveys became popular means of data collection. Although the awareness of biases resulting from the questioning process, several social scientists were interested in working on these applied methodological problems. This link between marketing and social sciences still exists today. The 1930s is the period of modern probability sampling approaches. Through the 1950s and 1960s, quantitative marketing research was the most used method, supported by digital computer era. The 1990s was characterised by technological computer advantages. The checkout scanners in supermarkets, computer assisted telephone interviewing, data analysis by microcomputer and remote terminals, interviewing through two-way cable television systems and internet surveys appeared. Therefore, new developments in information technology provided marketers with much richer information and tools.

The second half of the twentieth century has seen a dramatic increase in the number of numerical classification techniques available. As well as the increase in the variety of numerical classification methods, the expansion has taken place in the application fields. The quantitative model building approach moved forecasting into new areas: *Numerical taxonomy* is applied in biology, *Q analysis* in psychology, and *unsupervised pattern recognition* is commonly applied in Artificial Intelligence. *Clumping and grouping* is also applied in some fields. However, the more generic term is cluster analysis (Everitt, 1993).

Particularly in forecasting, significant gains have been made in marketing, especially since the 1960s. Advances have occurred in the development of qualitative and quantitative studies such as Delphi role playing. The quantitative model building approach moved forecasting into new areas. For instance, improvements based on statistical data, such extrapolation or econometric methods occurred. In the 1990s, gains have come from the integration of statistical and judgmental forecasts. Also the sophistication and design of new techniques is a clear feature of the 1990s and 2000s.

The 1990s and 2000s seems to be introducing a new stage for marketing research. Not only researchers are testing new techniques, but also they are applying techniques originated in another field to marketing. For example, using survival analysis in financial services (Thomas, 2000; Harrison and Ansell, 2002), neural networks in finance (Serrano-Cinca, 1996), and chaos theory in marketing (Hibbert and Wilkinson, 1994).

Hence, Artificial Intelligence techniques, initially developed in other fields, are increasingly being applied in business, particularly in finance and marketing. AI is becoming an alternative approach which deals with stored data and helps managers to achieve customer learning. A fuller description of the most applied AI techniques is explained in the next Chapter 4.

According to Hair, Bush and Ortinau (2003), five major trends in marketing research are commonly accepted by researchers and practitioners. The first is an increasing interest in secondary data collection methods. The second one is the movement toward technology-related data management. The third is an increased use of computers for information collection and analysis. The fourth one is related to the broader international database. Finally, the last trend anticipates the movement away from the pure data analysis towards data interpretation /information.

From a retailer point of view, all these improvements mean new challenges and new opportunities for managing information. Several retailers have been investing a lot of

money in data collection tools such as scanners, loyalty cards, customer service surveys, etc., in order to obtain as much information as they can about each customer and their relationship within the store. However, the main challenge appears when retailers have to look for the tools which enable them to analyse, interpret and shift this enormous quantity of data in interesting knowledge (Baron and Lock, 1995).

3.2 Forecasting as a marketing research task

As previously mentioned, marketing research is the process which provides information for enhancing decision-making processes in marketing. This information can be managed for deciding what to do and how to do it or for evaluating how effectively marketing is being done. As soon as the marketing research task endeavours to predict the future, various types of forecasting applications, models, methods or techniques can be employed.

'Forecasting is the prediction of values of a variable based on known past values of that variable or other related variables. Forecasts also may be based on expert judgments, which in turn are based on historical data and experience' (Makridakis, Wheelwright and Hyndman, 1998:599).

According to its definition, it is important to remark that forecasting is not planning. Although planning is about anticipating and organizing the various parts of the business to reach predetermined objectives, it is not the same as forecasting. According to Newman and Cullen (2002), planning involves the use of the resulting forecasts to help retail managers to make good decisions about the most attractive alternatives for the organization (in general terms), or in particular for the Marketing department, by for instance setting up the ranges of products or services customers will want in the coming years.

Forecasting is an integral part of the decision-making activities. The need of forecasting increases as managers attempt to decrease their dependence on uncertainty and change. Essentially, there is a wide range of forecasting problems within an organization. A distinction between two levels of forecasting labelled as

Macro forecasting level and Micro forecasting level is noted as follows:

Macro Forecasting Level

The fact that all the areas in a company are related, generates the need to develop a forecasting system which establishes a mutual relationship among forecasts made by different management areas. This system is quite complex to build because it considers different elements, as coming from the external environment as related to the areas within the organization.

Econometric models are likely to help when explaining complex forecasting systems. Econometrics, the interdisciplinary research involving statisticians and economists, has mainly been focused on measurement and prediction of the future economic activity (internationally, nationally or regionally). Constructing forecasting models to predict macroeconomic factors were the research interest of several economists (e.g. Klein and Goldberger, 1955). There are several econometric models. For instance, the Brookings model contains more than 200 equations built to test and evaluate the impact in the economic activity of the application of different economic policies (Makridakis and Wheelwright, 1978). A higher macroeconomic forecasting vision is found in the literature (e.g. Klein and Young, 1980; Adams, 1986; Granger, 1999). Although econometrics provides the tools for business to account for changes in the economy on their forecasts (Adams, 1986), they are mostly exploited by political and governmental issues. However, it is important to remark that some companies use econometrics to develop forecasting systems (Makridakis and Wheelwright, 1978).

Micro forecasting level, a Marketing interest

Despite the fact that the macro forecasting level intends not only to predict events which are relevant to more than one department of a company but also the environmental changes which can affect these events, it is important to note that each business department needs to find out specific predictions. Therefore research in forecasting is extensive in management. Moreover, it is known that the more uncertain the environment becomes, the more helpful forecasting is likely to be for

managers, in particular for their planning and execution tasks (Armstrong and Brodie, 1999). Table 3.1 illustrates some of the uses of specific forecasting into this micro forecasting level (Bolt, 1994: 64-65). Each use of forecasting is classified according the expected time horizon and its department interest.

Table 3.1 Uses of forecasting in some organization function areas

Time horizon	Marketing	Inventory	Finance	Purchasing	Top management
Short term (up to 3 months)	-Sales of each product type -Sales by geographical area -Sales by customer -Competition prices -Sales force targets	-Demand of each product -Demand for material -Demand for semi-finished products -Weather conditions	-Sales revenue -Production costs -Inventory costs -Leading indicators -Cash in-flows -Cash out flows	-Production -Cash availability -Purchasing of supplies and material	-Competition evaluation
	-Total sales Product categories -Major products -Product groups -Stock levels	-Demand for material	-Total demand -Inventory levels -Cash-flows -Short term - borrowing -Prices	-Demand for products -Demand for material -Lead time for purchasing	-Total sales -Sales breakdowns Prices
	-Total sales -Product Categories -Prices -General economic conditions -Promotional emphasis	-Total sales Expansion of warehouses	-Budget allocations -Cash flow	-Demand for products -Demand for raw and other materials	-Demand for sales -Other expenses -Cash-position -General economic conditions -Controls -Objectives
Long term (2 years or more)	-Total sales -Major product categories -New product introduction -Saturation points		-Total sales investment selections -Capital expenditure -Allocations -Cash-flows	-Contracts for buying of raw material	-Total sales -Costs and other expenses -Social and economic trends -Goals and objectives and strategies establishment -New products

Source: Adapted from Bolt, G. *Market and Sales Forecasting: A total Approach*. London: Kogan Page Limited, 1994:p64-65.

As far as the marketing department is concerned, the mentioned distinction between statistical forecasting and behavioural forecasting is again seen. Although the

majority of the forecasting needs listed in the second column of Table 3.1 are focused on sales, it is important to note that there are other types of forecasting such as new product introduction (product choice) and promotional aspects (targeting actions). In both cases, knowing customers' future behaviour is the basis to take proper decisions. Knowing customers' future behaviour is not an easy task. Customers are not simple robots. As mentioned in Chapter 2, they may behave differently according to the personal situation or particular circumstances. Related to that, it is also important to consider the fact that customers do not always know for sure what they are going to do into the near future. Therefore, knowing their attitudes or intentions related to a new product introduction, brand choice or store choice is not always enough, in managerial terms (See Section 3.4).

Despite the fact that is difficult to list all the large number of forecasting approaches available to firms today, particularly in the food retail industry, we would like to note the academic and managerial interest in sales forecasting (e.g. Venugopal and Baets, 1994; Huth, Epright and Taube, 1994; Morwitz, Steckel and Gupta, 1997; Lee, Elango and Schnaars, 1997; Frees and Miller, 2004), layout and merchandising (Newman, Yu and Oulton, 2002), new products and services (Jamieson and Bass, 1989), market share (Brodie and de Kluyver, 1987; Agrawal and Schorling, 1996; Klapper and Herwartz, 2000), customer purchasing intentions (Kalwani and Silk, 1982) and customer behaviour (Davies, Foxall and Pallister, 2002; Buckinx and Van den Poel, 2005).

One of the interests of this thesis is to provide an overview the main forecasting approaches to forecast customer behaviour. Therefore, from now on, and in order to limit the wide scope of this area, a particular mention to this micro forecasting level will be assumed, and predominantly related to forecasting customer behaviour in the food retailing industry. Based on that, the description of the main stages required when forecasting are explained in the next section.

3.3 The forecasting process

For the purpose of this research, Armstrong's forecasting process (2001) will be followed as the main reference within this thesis. The author provides a checklist of principles designed to assist forecasters and managers to systematically evaluate the forecasting processes they use to increase accuracy and reduce costs. These principles, cover the sequential stages of a forecasting process which include:

- a) formulating the problem,
- b) obtaining information,
- c) selecting and implementing forecasting methods,
- d) evaluation of forecasting methods,
- e) using forecasts.

3.3.1 Formulating the problem

Determining the purpose and objective of the forecast is the first step of the process. The forecasting research objective needs to be understood to provide the information that will facilitate the decision-making process (Bails and Peppers, 1993). A suggestion is provided by Georgoff and Murdick (1986) when formulating the problem which consists of answering questions organized into four major categories: time, resource requirements, input and output. In reference to time, the purpose of forecast needs to dictate the length of the forecast period. When making predictions about the near future, the forecasts are called short term (3 months approximately); when considering the very distant future, they are called long term (2 years), and the intermediate case involves middle term forecasts. It is important to note that different forecast horizons normally require different approaches. Resource requirement, input and output is related to the information and its accessibility.

3.3.2 Obtaining information

Once the problem is formulated, a selecting process focused on how and where to

obtain the necessary data to solve the problem is essential. Sometimes the existence of sufficient historical information is one of the main determinants of the subsequent steps in the forecasting process. However, when there is an existing lack of data, companies and researchers need to look for it. Data collection procedures have a prominent part to play in management (Kotler, 2000). The information is gathered through different approaches including internal company databases, external information systems and other types of marketing research. The aim of this section is to review the main data collection methods, according to the nature of data (primary or secondary) and the type of research (qualitative or quantitative). Table 3.2 summarises the main data collection sources.

Table 3.2 Data collection sources

Types of data collection methods	
PRIMARY DATA	SECONDARY DATA
<p>Qualitative collecting methods</p> <ul style="list-style-type: none"> - Exploratory: deep interviews, experts opinions interviews, focus groups - Observational: case-based research, ethnography, grounded theory, others <p>Quantitative collecting methods</p> <ul style="list-style-type: none"> - Surveys - Behavioural patterns observations 	<p>Database collecting methods</p> <ul style="list-style-type: none"> - External Data: periodicals and books, Internet websites, commercial data, government publications and prior research reports, others - Internal Data: scanners, loyalty cards, coupons redemptions, customer's support service information, large product, prices database

Primary data

In reference to primary data, Kotler (2000:106) states that primary data

'are data gathered for a specific purpose or for a specific research project'.

Primary data can be collected using different methods (See Table 3.2). Some of them are based on qualitative research and others are based on quantitative research.

Qualitative data collection methods

The prior objective of qualitative research methods is to achieve a broader understanding of the consumer and to gain insights into their behaviour, motivations, preferences, attitudes and needs. Malhotra and Birks (2000:156) define qualitative research as

'an unstructured, primarily exploratory methodology based on small samples, intended to provide insight and understanding'.

Six types of qualitative collection methods are commonly used in marketing research which are split into exploratory (e.g. in depth interviews and focus groups) and observational (e.g. case-based research, ethnography, and grounded theory).

Related to the exploratory methods, *individual in-depth interviews* are those that take place face to face between the researcher and the consumer. Depending on the data required, either consumers will be interviewed or experts in the field, such as executives, technical experts and thought leaders. The objective of this type of interview is to achieve insight into the market place directly from chosen candidates. This allows for an initial understanding of the consumer, allowing the researcher to begin to discern the ranges of attitudes consumers have to a certain subject in a relatively short period of time. There are two types of in-depth interview methods: semi structured and non-directive. The difference between the two methods lies in the amount of guidance the interviewer provides. In semi structured interviews, the interviewer covers a specific list of topics whereas the non-directive interviews are much more open ended, and thus usually take more time than the semi structured interviews (Calder, 1994). These interviews are useful for researching key product benefits, to trigger creative insights into existing products and to gather basic market insights such as trends in technology and market demand. The main advantage of individual interviews is that it is easier for the researcher to control the time dedicated to each issue than it is in focus groups and also shyness or reservations on behalf of the interviewee may be easier to overcome in a one-to-one setting. Sometimes, the interview is answered by an expert or a group of experts. *Expert*

opinions interviews are very commonly applied in forecasting. According to Ashton (1986) the number of experts ranges from 5 to 20. The experts are asked to predict the behaviour of a specific issue (e.g. market, consumers, demand or trend). Delphi procedures is an iterative structured survey which experts make forecasts for a problem, receive anonymous summary feedback on the forecasts made by other experts, and then make a further forecast.

As far as focus-groups are concerned, focus groups are discussions which consist of dynamic group interviews, where a group of consumers or experts are gathered together in order to communicate insights into the area in which the organization is interested, the key characteristic of this method is the emphasis put on interaction made between participants. The main purpose of the focus group is not listening to people's opinion but to

'gain insights by creating a forum where respondents feel sufficiently relaxed to reflect and portray their feelings and behavior- using their language and logic' (Malhotra and Birks, 2000:161).

Compared to in depth interviews, focus groups allow for more stimulation as there is more participation from all sides, and often the presence of a group encourages people to speak out (Carson et al., 2001). Exploratory focus groups are frequently used in the preliminary stages to a survey. For this reason, this type of focus group is considered to be limited, to either testing survey questions or generating, but not validating insights (Calder, 1994). One of the purposes of these exploratory groups is to obtain insights into a new area through the generation of new ideas. Once these focus groups have been completed, further marketing research is required to assess how many people think an idea is a good one and to answer other questions about the ideas. Although focus groups can be very powerful, Armstrong and Brodie (1999) strongly suggest not using this method in forecasting, in particular when a Delphi procedure is carried out. The reason is that experts' forecasts should generally be independent of one another.

The observational methods differ from the exploratory techniques as they do not expect verbal answers but generate data by watching ways of behaviour. The purpose of *case-based research* is not to represent the world, but to represent the case. Criteria for conducting the kind of research that leads to valid generalization need modification to fit the search for effective particularization. The utility of case research to practitioners and policy makers is in its extension of experience. The methods of qualitative case study are largely the methods of disciplining personal and particularized experience (Stake, 1994). A type of case-based research is *role playing* which simulates a realistic business situation. Role playing is useful for making forecasts of the behaviour of individuals who are interacting with others, and especially when there is a high risk of conflict. Therefore, a person's role may be a dominant factor as in predicting how someone in a firm would behave (Armstrong and Brodie, 1999).

Ethnography is originally developed by anthropologists who spend significant periods of time with local people making detailed observations of their practices and attempting to understand their culture (Hammersley and Atkinson, 1983). This practice is now being transferred into marketing research in order for marketers to uncover emerging and unmet needs of the consumer. While traditional ethnographers tend to immerse themselves in cultures for weeks or months, marketing researchers are limited according to their time schedule (Carson et al., 2001). A key characteristic of ethnographic research is that it takes place in realistic scenarios. An example of a retailing ethnographic study is provided by Dunnet and Arnold (1999) when they explain Jane's experience when she decided to work for Wal-Mart, in order to investigate how the strong organizational culture of the retailer impacted upon customer satisfaction.

Grounded theory has emerged as one of the most popular and rigorous methods of deriving theories from qualitative data in sociology. However, its use in marketing is limited (Carson et al., 2001). A guide to handling masses of qualitative data and systematically analysing them is provided by Glaser and Strauss (1967) and Strauss and Corbin (1990; 1997).

Due to the dynamism of the marketplace and the necessity of business to be competitive, new methods of research are constantly increasing. Beyond the traditional data collected from qualitative methods, visual data are becoming very important. The use of video and film is being used in different areas (Denzin, 2000). There are some companies which are using different qualitative methods to understanding phenomena and to gain meaningful insights into certain consumer circumstances and changes. For instance, IDEO presents 51 qualitative research methods used to discover individual insights and help businesses to inspire design. Five of these methods are card sorting, word concept association, role playing, cognitive maps and conceptual frameworks¹.

Before continuing, it is important to mention that not always is access to certain information easy. For instance, as shown in Figure 3.1, the access to direct answers might be difficult due to the complexity of the subject or due to the fact that the subject is considered private or difficult to discuss. Unrecognized needs are difficult to access due to the unconsciousness of their nature. Therefore, if the information required of the consumer is considered by the consumer to be private, to be non communicable and the consumer is even unaware of it, then the layers of response from the consumer will most likely be unconscious (Malhotra and Birks, 2000). If the consumer considers all the above but is aware of the need, he/she has then the responses he/she will provide will be more intuitive and imaginative. As the information required of the consumer becomes more public and the consumer becomes more able to communicate his/her needs, then the responses given will be more and more reasoned. The following table illustrates this:

¹ A detailed description about these methods is found in <http://www.ideo.com>

Figure 3.1 Relationship between expected consumer responses and research method designs used

Access to respondents			
Private			Public
Non communicable		Communicable	
Unaware	Aware		
Layers of response from respondents			
Unconscious Repressed	Intuitive Imaginative	Concealed Personal	Spontaneous Reasoned Conventional
Tends to be Exploratory Research			Tends to be Conclusive Research

Source: Malhotra, N. K. and Birks, D. F. *Marketing Research: An Applied Approach*. UK: Prentice Hall, 2000:p74

Quantitative Data Collection Methods

According to Malhotra and Birks (2000:156) definition,

'quantitative research is a research methodology that seeks to quantify the data and, typically, applies some form of statistical analysis'.

The objective of quantitative research is to achieve a quantifiable sample of consumers, which enable the decision makers to extrapolate it statistically, and, thus, obtain close estimations on consumer attitudes, intentions and/or behaviours. The most common quantitative primary data collection methods are both surveys and observational studies. A *Survey* consists of

'a structured questionnaire given to a sample of a population and designed to elicit specific information from respondents' (Malhotra and Birks, 2000:209).

There are several types of survey methods, including personal interviewing, mail surveys, fax and web surveys, self-administered questionnaires, omnibus surveys,

telephone interviews, and computer interactive interviewing. There is a set of factors affecting the choice of a survey. These factors include the research objective, the type of population to be interviewed, the type of data required, the question content, budget, available facilities, time and ethical issues. Particularly in forecasting, intention surveys are focused on asking people how they would behave in various situations. They are often used when sales data are not available, such as for new product forecasts.

'Observation is based on watching what people do, looking at their behavioural patterns and actions and at objects, occurrences, events and interactions' (Carson et al., 2001:133).

Observation is a very useful research method when patterns of behaviour are analysed, but it is less useful in respect of information on consumer intentions. Customers and shoppers behaviour can be observed in a store. Contrarily to behavioural data (internal databases) which track each customer's real behaviour, these observations help retailers to have a better knowledge about what happens in their shops, but without tracking each customer. In addition, observation methods collect information which, currently, it is difficult to collect from scanner systems. For instance, in general, scanner systems and loyalty card programmes do not capture how much time a customer spends in a store, or how many products he/she consulted before selecting one or which sequence of product categories was followed before purchase. Nor is a survey appropriate in this case because customers do not know the exact answer. Therefore, retailers who want to be aware of this type of information need to implement human observations or mechanical devices (video cameras, eye cameras, audiometer, or web cams) which record all this kind of information from the stores. Grove and Fisk (1992) provide more information about these observation methods.

Secondary data

According to Kotler (2000:106) secondary data can be described as data that

'were collected for another purpose and already exist

somewhere'.

Sources of secondary data are mainly split into external and internal data. Secondary external databases comprise published or existing material such as government statistics, economic data, trade information published surveys (Bolt, 1994) and panels (consumer data, retailer panels and audience panels). INE (Instituto Nacional de Estadística) and ICEX (Instituto Español de Comercio Exterior) are the main government statistics sources relevant to this thesis. Information about census, economic data and social trends are accessible in these sources. In addition market reports are also provided by several companies. Both Spanish and international market reports are accessible for example at EUROMONITOR, ALIMARKET, GUIAME, DATAMONITOR, MINTEL. Related to panel data, TNS, IRI and EGM are the leaders. Updated information databases may be found also online such as PROQUEST, HARVEST, MAGIC and GUIAME. Both qualitative and quantitative information is collected from external databases, depending on the source.

The advantages of the use of secondary data are that it is essentially cheap to find (compared to primary data) and also relatively fast. In addition, secondary data often provides information about the environment and the marketplace. However, to be really useful, existing information should be regularly updated. Normally, secondary data sources are consulted prior to the design of some primary data collection method.

Related to the internal data, there are quantitative data collection tools including scanners, catalogue purchase records, loyalty card programs and specific internal customers, sales, channels, prices and product databases which offer an attractive opportunity for retailers. According to Baron and Lock (1995), there is scope for efficiency savings and exploitable competitive advantage when translating this enormous volume of data into meaningful information. These tools enable access to a type of information which previously were provided only by indirect marketing research survey and panels. But there are some limitations. According to Kotler (2000), the tools such as scanner systems and loyalty card programs which are able

to collect and warehouse all the traces customers leave from their purchases are called behavioural data collection tools. Despite the fact that tracking customer behaviour is likely to be an advantage, often, the volume of the raw data generated in the superstore is so huge that it poses problems even for simple analysis (Baron and Lock, 1995). In these cases, the development of new methods is required.

'These methods need to include pre-specified partially or wholly automated analyses and knowledge-based or AI systems to make sense of large volumes of potential output and to reduce it to manageable volumes' (Baron and Lock, 1995:51).

As the experiments reported in Chapter 5 and Chapter 6 of this thesis shows, internal data collected from the loyalty card program and scanner systems of a particular supermarket chain can be used to forecast patterns of customer behaviour through an AI approach². According to Mitchell, Russo and Wittink (1991), there is a variety of approaches to analyse scanner data which includes human judgement, expert systems and statistical models. A fuzzy learning technique is used in our research (See Section 4.6).

3.3.3 Selecting and implementing forecasting methods

Based on whether or not there is sufficient historical information available, the majority of researchers within the literature clearly distinguish between quantitative/statistical forecasting methods and qualitative/judgmental forecasting methods, depending on whether the analysis of the raw data is based on quantitative methods or qualitative methods. Normally, when a large quantity of data is available, quantitative research methods are applied. Moreover, when there are few records, the analysis tends to be qualitative. However, there is not a specific rule for that. Based on that, two main increasing trends in forecasting are found over the literature. The

² Despite the fact that traditionally database information have been considered a type of secondary data, we would like to note that definitions of primary data and secondary data may be somehow confusing. Primary data relates to all the data gathered for a specific purpose. The data applied in the experiments, already existed in the internal database. However, a selection of some variables was specially extracted for each segments. Then, although obtaining some specific information could be considered primary data, the traditional way is assumed in this thesis. Based on that, behavioural data is considered to be a type of secondary data.

first one is based on combining different methodologies (Clemen, 1989; Blattberg and Hoch, 1990; Bunn and Wright, 1991; Batchelor and Dua, 1995; Armstrong, 2001). The second attempts to improve and make more sophisticated the existing techniques (Aaker and Jacobson, 1987). A review of the most common applied research methods when forecasting are summarised in Table 3.3.

Table 3.3 Types of forecasting research methods

QUALITATIVE METHODS	QUANTITATIVE METHODS
<ul style="list-style-type: none"> - Judgemental bootstrapping - Empirical reports - Analogies, extrapolation 	<ul style="list-style-type: none"> - Time series models: Exponential smoothing, decomposition, autoregressive/moving average (ARMA, ARIMA), multivariate autoregressive models and State space models or Kalman filter models - Causal models: Simple regression, multiple linear regression, econometrics, cross-sectional regression, dynamic regression models, intervention analysis (special case of dynamic regression models)

Qualitative forecasting methods

According to Armstrong and Brodie (1999), qualitative forecasting techniques can be applied when little or no quantitative information is available but sufficient qualitative knowledge exists. They are mainly based on qualitative and judgmental data.

Judgmental bootstrapping approaches convert subjective judgments into objective procedures. The interviewed opinions are converted to a set of rules. Then, the developed model is used to make following forecasts.

Empirical reports are non-numerical reports based on the qualitative data gathered from the interviews, focus groups or other qualitative collecting methods which summarise both the observed behaviour and deal with personal opinion. Usually, all the data collected by qualitative methods are analysed and edited by methods originally created to analyse text. However, the methods of analysis have been

improved to be able to deal with different sort of data such as text, sound and video (Flick, 1998).

Subjective extrapolation from past situations of results from these situations can be used to predict the situation of interest (Duncan, Gore and Szczypula, 2001). In most of the cases, experts are asked to identify analogous situations (Armstrong and Brodie, 1999).

Quantitative forecasting methods

A detailed description of the quantitative methods is provided by Makridakis, Wheelwright and Hyndman (1998). Quantitative forecasting techniques can be applied when there is sufficient historical information available and it can be quantified in some form of numerical data. Moreover, there is an assumption that some aspects from the past will continue in the future (Makridakis and Wheelright, 1989). There is a wide set of forecasting techniques, but researchers agree when distinguishing between two main groups: Time series models and causal (explanatory) models.

In particular, a time series model assumes that some pattern or combination of patterns is recurring over time. This type of model does not search for explanations but to identify how the system is going to evolve in time, accordingly to the past. Therefore, the method does not assume that the variable to be forecasted exhibits an explanatory relationship with one or more independent variable. According to Makridakis and Wheelwright (1978), time series model treats the system as a 'black box' and makes no attempt to discover the factors affecting its behaviour. Accordingly, when there is no intention to know why something happens, just to predict what will happen, times series seem to be the most useful quantitative technique. Exponential smoothing, decomposition, autoregressive/moving average (ARMA, ARIMA), multivariate autoregressive models and state space models (or Kalman filter models) are some of the most common time series methods in management.

Differing from time series, explanatory methods show and explain the relationship between the inputs and outputs. Basically, explanatory methods assume that the value of a certain variable (output) is a function of one or more other variables (inputs).

'Developing an explanatory model facilitates a better understanding of the situation and allows experimentation with different combinations of inputs to study their effects on forecasts' (Makridakis and Wheelwright, 1978: 159).

Simple regression, multiple linear regression, econometrics, cross-sectional regression, dynamic regression models, intervention analysis (special case of dynamic regression models) are the most common techniques within this group³.

3.3.4 Evaluation of forecasting methods

'From a predictive point of view, the evidence is equally clear that there is no way of knowing a priori who or which method will be more accurate' (Makridakis and Wheelwright, 1989: 5).

To evaluate the forecasting accuracy of each method, some authors propose a comparison between methods (e.g. Armstrong, 1985; Brodie and de Kluyver, 1987; Makridakis, 1993; Kumar, Rao and Soni, 1995; Ansuj et al., 1996; Agrawal and Schorling, 1996; West, Brockett and Golden, 1997). Benchmarking is also suggested by other researchers (e.g. Mentzer, Bienstock, and Kahn, 1999).

Qualitative versus quantitative approaches

Despite the fact that there are some general statements in forecasting research, such as the suggestion of applying a qualitative forecasting model when predicting long term forecasts or the necessity of working with similar data to the past patterns and past interrelationships when extrapolating the future, there is no accepted group of rules which indicates the most appropriate forecasting model.

³ A brief review of the forecasting models is offered in Lilien and Kotler (1983). More detailed descriptions are provided in forecasting textbooks such as Makridakis, Wheelwright and Hyndman (1998).

In fact, both groups, qualitative and quantitative approaches are not autonomous or incompatible. Conversely, the relationship between them has been demonstrated to be increasingly powerful in forecasting research. According to Armstrong and Brodie (1999), there is a rising amount of integration between judgmental and statistical data, which is used to improve forecast accuracy (e.g. Clemen, 1989; Godwin and Wright, 1993; Batchelor and Dua, 1995; Armstrong and Collopy, 1998; Armstrong, 2001).

For example, *Rule based forecasting* is a type of expert system that allows integration of managers' knowledge about the domain with time series data in a structured and inexpensive way. Then, in many cases a useful guideline is that trends should be extrapolated only when they agree with managers' prior expectations. Furthermore, as the name implies, *expert systems* use the rules of experts. These rules are typically created from protocols, whereby the forecaster talks about what is being done while making forecasts. The real promise, however, is for expert systems to draw upon empirical results of relationships that come from econometric studies. In fact, this is a common way to construct expert systems. Expert opinion, conjoint analysis and bootstrapping can also aid in the development of expert systems (Armstrong and Brodie, 1999).

Nevertheless, there are still some researchers who attempt to demonstrate the superior quantitative models accuracy and the opposite way round. For example, Lee, Elango and Schnaars (1997) demonstrated that time series extrapolations provided more accurate forecasts than the judgmental approach when predicting buying intentions for consumer durable goods. On the other hand, Armstrong, Morwitz and Kumar (2000) suggested that purchase intentions provided better forecasts than a simple extrapolation of past sales.

It is clear that research in forecasting is extensive. Not only it includes many forecasting methods but also combinations between them. Despite the fact that there are several examples of forecasting research in business and several studies to

evaluate which is the most precise, no ranking of the most effective and efficient forecasting method exists within the literature. Consequently, Makridakis and Wheelwright's point of view is completely shared within this chapter, which is the following:

'There is no published evidence that any individual forecaster or forecasting method has been consistently and significantly more accurate than any other'. It is clear that some are more successful than others, at least for short periods of time. (...) However, from a predictive point of view, the evidence is equally clear that there is no way of knowing a priori who or which method will be more accurate (Makridakis and Weelwright, 1989: 5).

However, it is important to note that there are several variables which may affect the accuracy of the results. Table 3.4 lists some of the variables which may influence to evaluate the forecasting technique, such as time, cost, accuracy, etc.

Table 3.4 List of possible criteria of evaluating forecasting

CRITERIA
<ul style="list-style-type: none"> - Time (short, medium or long term) - Cost (data accessibility and resources) - Accuracy - Realism - Assessing Uncertainty - Forecasting proposal - Technical sophistication - Variability and consistency of data - Amount of detail necessary - Turning points (opportunity or caution moments)

Source: Adapted from Georgoff, D.A. and Murdick, R.G. "Manager's guide to forecasting." *Harvard Business Review*, 1986, 64 (Jan-Feb), 110-120.

3.3.5 Advanced forecasting techniques: A general overview

Apart from the research tendency of combining different techniques in order to improve forecasting results, there is also a growing trend focused on developing new forecasting techniques. Recently, forecasters have extensively examined complex

procedures oriented to increasing the sophistication of previous models (e.g. Aaker and Jacobson, 1987; Turksen and Willson, 1995), or by designing completely new advanced and mixed techniques. In reference to the design of new advanced techniques, the quantitative model forecasting approach moved forecasting into new areas such as artificial intelligence field⁴ (See Chapter 4).

Statistical forecasting versus Artificial Intelligence forecasting

Despite the fact that there is a trend which combines statistical and artificial intelligence techniques when forecasting (e.g. Tseng and Tzeng, 2002), there is still a diverse opinion whether AI is useful or not. Several articles aim to provide a comparison between artificial intelligence approaches and statistical approaches (e.g. Montgomery, Swinnen and Vanhoof, 1997). This research line has been increasing during the last 20 years. For instance, Zhao, Collopy and Kennedy (2003) review the literature comparing artificial neural networks (ANN) and statistical models, particularly in regression-based forecasting, time series forecasting, and decision making with the intention to give a balanced assessment of ANN's potentiality for forecasting. It is important to note that although not in an extent way, some comparisons have been applied in the marketing field as well (e.g. Venugopal and Baets, 1994; Kumar, Rao and Soni, 1995; Agrawal and Schorling, 1996; Church and Curram, 1996; Ainscough and Aronson, 1999; Alon, Qi and Sadowski, 2001). Most of researchers agree that the primary difference between the approaches is when processing data. According to Venugopal and Baets (1994), neural networks are powerful alternative tools and a complement to statistics when data are multivariate with a high degree of interdependence between factors, when the data are noisy or incomplete or when there are many hypotheses to test. Koslowsky (2003) also states that sometimes neural networks are able to outperform a regression model, but many times they can not.

Not all specialists in the field believe in the complementarities between statistical

⁴ Noting that when referring to AI, often we refer to two AI subfields, which are learning and fuzzy logic, in particular neural networks and fuzzy techniques.

and artificial intelligence procedures when forecasting. On the one hand, there are researchers who defend AI techniques capabilities. According to Kuo, Wu and Wang (2002), statistical models such as regression are not feasible when the data is not seasonal, for instance in promotions. Therefore, Kuo, Wu and Wang (2002) consider artificial intelligence techniques much better than statistical procedures. The superiority of ANN networks compared to conventional statistical methods are also stated by Kumar, Rao and Soni (1995), Turksen and Willson (1995), Agrawal and Schorling (1996), Ali and Rao (2001), Tseng and Tseng (2002), Boone and Roehm (2002) and Chiang, Zhang and Zhou (2004). Also fuzzy logic is a field for improving marketing forecasts (e.g. Tseng and Tzeng, 2002; Kuo, Wu and Wang, 2002).

On the other hand, there are researchers who believe AI approaches are still at an immature stage. Neural networks have not produced more accurate forecasts than other methods (Chatfield, 1993). Although some promising work has been done in the study of market share forecasting (Agrawal and Schorling, 1996) and in predicting consumer's choice (West, Brockett and Golden, 1997), Armstrong and Brodie's (1999) advice is to ignore neural nets. According to Armstrong and Brodie (1999), the application of the more complex and sophisticated technique does not necessarily mean that they are going to be more useful. Table 3.5 summarises the strength of the different opinions found over the literature.

Table 3.5 Table of main contributions of the application of neural networks

Main Contributions of NN	Total
Neural Networks yield better results than other techniques when compared to them	45
Neural Networks yield similar results than other techniques when compared to them	14
Neural Networks yield worse results than other techniques when compared to them	3
One type of Neural Networks outperforms the others	18
An integrated/hybrid model including Neural Networks improves the results of the study	21
Neural Networks are deemed for promising for future developments of the application	26
Neural Networks are shown to offer new insights into the application	29
Neural Networks are utilised in a 'real world' case of the application	9

Source: Vellido, A. Lisboa, P.J.G. and Vaughan, J. "Neural networks in business: a survey of applications (1992-1998)." *Expert Systems with Applications*, 1999 17: p62

The higher number of publications collected by Vellido, Lisboa and Vaughan (1999)

shows that NNs perform better than other techniques. Also a considerable number of researchers agree with the opportunity to apply a hybrid model when applying these techniques in business, although the number of real applications is still low in the business field (See Table 3.5). According to the authors, there are generally fewer researchers who do not believe in the power of these AI techniques.

Moreover, the advantages and disadvantages of applying neural networks are listed in Table 3.6 (Vellido, Lisboa and Vaughan, 1999:62). Although a list of applications in marketing is provided in Section 4.5, at this point it is important to point out that the number of contributions enhancing the advantages of these techniques is higher than the number which demonstrates their disadvantages.

Table 3.6 The most frequently quoted advantages and disadvantages of NN

Advantages of NN	Total	Disadvantages	Total
NN are able to learn any complex non-linear mapping/approximate any continuous function	31	NN lack of theoretical background concerning to explanatory capabilities/NN as black box	28
As non-parametric methods, NNs do not make a priori assumptions about the distribution of the data/input-output mapping function	30	The selection of the Network topology and its parameters lacks theoretical background/ It is still a trial and error matter	21
NNs are very flexible with respect to incomplete, missing and noisy data/NNs are 'fault tolerant'	29	NN learning process can be very time consuming	11
NN models can be easily updated/are suitable for dynamic environments	15	NN can overfit the training data, becoming useless in terms of generalization	10
NNs overcome some limitations of other statistical methods, while generalizing them	15	There is no explicit set of rules to select a suitable NN paradigm/learning algorithm	8
Hidden nodes in feed-forward supervised NN models can be regarded as latent/unobservable variables	5	NNs are too dependent on the quality/amount of data available	6
NN can be implemented in parallel hardware, increasing their accuracy and learning speed	4	NNx can get stuck in local minima/narrow valleys during the training process	5
NN performance can be highly automated minimising human involvement	4	NN are still rapidly evolving and are still not reliable/robust enough yet	3
NN are specially suited to tackle problems in non conservative domains	3	NN lack classical statistical properties; Confidence intervals and hypothesis testing are not available	2

Source: Vellido, A. Lisboa, P.J.G. and Vaughan, J. "Neural networks in business: a survey of applications (1992-1998)." *Expert Systems with Applications*, 1999 17 51-70.

3.4 Forecasting customer behaviour

Forecasting customer behaviour has been a common shared challenge by academics and practitioners. As mentioned in Chapter 2, the majority of research in consumer behaviour was oriented to the individual/consumer, as the subject of research. Then, finding out his/her needs, attitudes, intentions and beliefs was one of the main focus of research. Some researchers considered that finding out the reasons why people shop would help to predict their behaviour. Maslow (1954) attempts to explain people's motives for shopping by introducing the pyramid of Maslow, a theory based on hierarchical human needs.

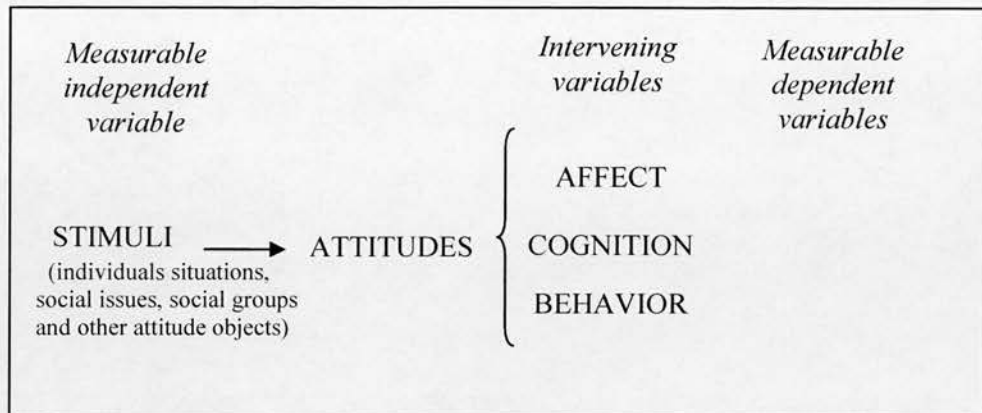
According to Maslow, there is an evolution of human needs, from the basics to the top of the pyramid. In order to reach the top it is necessary to follow up all the previous stages. Consequently, researchers based on Maslow's theory consider that the need to purchase products is the main reason for shopping. Since Maslow's pyramid was developed, several improvements, critics and alternatives have been appeared. McGoldrick (2002) provides an interesting historical evolution of these theories in Chapter 3 of his book.

Apart from attempting to predict customer behaviour by understanding their needs, there is also a research line aimed at demonstrating which are the most suitable measures to forecast customer behaviour. This research perspective is generally based on the schematic conception of attitude (Rosenberg and Hovland, 1960).

The authors conceptualized the attitude as a multicomponent representation of three components. As figure 3.2 illustrates, all the responses to a stimulus-object are mediated by the person's attitude toward that object. Based on Rosenberg and Hovland's representation of attitude, Fishbein and Ajzen (1975:340) state

'attitudes are viewed as a complex system comprising the person's beliefs about the object (cognition), his feelings towards and object (affect) and his action tendencies with respect to the object (behavior)'.

Figure 3.2 Schematic conceptions of attitudes



Source: Adapted from Rosenberg, M.J and Hovland, C.I. Cognitive, affective and behavioural components of attitudes. In C.I. Hovland and M.J. Rosenberg (Eds.), *Attitude Organization and Change: An analysis of consistency among attitude components*, New Haven, Conn: Yale University Press, 1960.

According to the definition, it seems that in order to forecast the human response (behaviour) to a stimulus, beliefs, affects and behaviours are likely to be identified. In fact, whether there is a clear link between attitudes and actual behaviour has been a priority by psychologists during the last 50 years (Ajzen and Fishbein, 1977; Bentler and Speckart, 1979; Ajzen and Fishbein, 1980). Some authors were also interested in finding out whether there was a positive correlation between intentions and actual behaviour. According to Fishbein and Ajzen (1975), the best predictor of an individual's behaviour will be a measure of his intention to perform that behaviour.

Moving the attitude-behaviour relationship from social psychology to the marketing field, several points of view appear. Not only is there not an evident conformity about the most suitable predictor to forecast behaviour, but also there are a wide range of different statements and perspectives, mostly depending on the situation.

A number of publications support the existing relationship between attitudes and behaviours (e.g. Bagozzi, 1981; Allen, Iacono and Danielson, 1992; Baldinger and Rubison, 1996). Others consider that having a positive attitude towards an object does not necessarily mean that the

'predisposition to respond to that object in a favourable manner'
(Fishbein and Ajzen, 1975:222).

the correlation between both attitude and the coherent behaviour happen (Kamakura and Gessner, 1986). Apart from these two extreme perspectives, there is a debate attempting to demonstrate whether the behaviour is determined by attitude as a whole or merely by some of its components (Fishbein and Ajzen, 1975). For instance, Madrigal (2001) firstly suggested the cognition element of the attitude as the most predictive measure. According to Zaltman (2003), approximately 95% of all cognition occurs below awareness in the shadows of the mind while only 5% occurs in consciousness. Others consider emotions as primary motivations of behaviour (Holbrook and Hirschman, 1982; Zanna and Rempel, 1988; Cohen, 1990; Allen, Iacono and Danielson, 1992).

Consequently, depending on the situation, it is demonstrated that certain behaviours respond most to the cognitive component, others to the affective component and still others to the conative component.

Several publications in the literature aim at finding whether attitudes predict purchasing intentions (Madrigal, 2001). However, in marketing research, there is an evident interest to finding out whether purchase intentions are predictive measures of purchase behaviour (Juster, 1966; McNeil, 1974; Tauber, 1975; Morrison, 1979; Warshaw, 1980; Infosino, 1986; Jamieson and Bass, 1989; Mansky, 1990; Armstrong, Morwitz and Kumar, 2000). Most of them concluded there was a positive link between intentions and behaviours. However, not always a strong link was easy to demonstrate. For instance, Belk (1985) reviewed the factors that may explain discrepancy between stated purchase intentions and subsequent behaviour. According to Fishbein and Ajzen (1975), 3 main requirements (specificity, stability and volitional control) need to be accomplished whether a perfect correspondence between intention and behaviour is going to be achieved. Moreover, this factor also helps to explain not only the reasons why an intention does not become behaviour but also why a non-intention becomes behaviour.

Some authors agree with the utility of intentions data to forecast behaviour but under certain conditions (Armstrong, 1985; Morwitz, Steckel and Gupta, 1997). For instance, Morwitz and Schmittlein (1992) state that the shorter is the forecasting horizon, the more valid the purchasing intentions data are to forecast behaviour. Other publications state that purchasing intentions may be accurate to predict behaviour, depending on the product/service planned to purchase (e.g. Adams, 1986; McNeil, 1974; Tauber, 1975; Silk and Urban, 1978; Warshaw, 1980; Kalwani and Silk, 1982; Infosino, 1986; McQuarrie, 1988; Jamieson and Bass, 1989; Morwitz and Schmittlein, 1992). Apart from the product intended to be purchased, many researches insist that the factor which influences the predictive degree of intentions is the people. Related to this some authors state past experience with the real behaviour affects the relationship between intention and later behaviour (Fazio and Zanna, 1981; Kalwani and Silk, 1982; Ronis, Yates and Kirscht, 1989; Morwitz and Schmittlein, 1992; Bemmaor, 1995). Consequently, this explains that sometimes a segmentation analysis previous to the forecasting process is required (Bolt, 1994). Investigations on whether or not the use of segmentation can improve the accuracy of sales forecasts based on stated purchase intention are provided by Morwitz and Schmittlein (1992). Also Venugopal and Baets (1994) use segmentation approaches prior to retail sales forecasting.

Conversely, declarations that intentions fail to predict behaviour are also present in the literature (Davies, Foxall and Pallister, 2002). To avoid imperfect correspondence between verbal intention statements and real behaviour, authors not only believe in observing behavioural criteria to predict behaviour but also suggest abandoning the intention-behaviour hypothesis, which lies at the center of psychological and marketing theorizing.

As mentioned in Chapter 2, the aim of this thesis is to forecast customer behaviour from a retailer perspective. From our point of view, the subject of research is customer's behaviour, and particularly the household shopping behaviour data. Based on that, primary data related to needs, attitudes or intentions of the customers are not going to be used in the empirical research. Secondary data coming from

internal data bases are the input for our experiments. Therefore, past customers' behavioural data collected from both scanner systems and loyalty card program are going to be used to forecast future customer's behaviour.

3.4.1 Behavioural data

According to Fishbein (1963) there are three main types of behavioural criteria: Single observation of a single act, repeated observations of the same single act and multiple act criterions, which is based on single or repeated observations of different behaviours.

As is illustrated in Figure 3.3, the basic element of the scheme is the single act criterion. It consists of 4 elements:

- behaviour,
- target,
- situation and
- time.

In contrast to the single act, the other two types vary in terms of their specificity with respect to at least one element of the behaviour. The second type can be obtained by observing the same specific behaviour directed at different targets, in different situations or at different times. Multiple act data is obtained by computing some index across observations of different behaviours. It is shown that multiple act criterions are based on observation of several behaviours with respect to a given target, in a given situation at a similar period of time.

Figure 3.3 Types of Behavioural Criteria

		Observations							
		1	2	3	...	j	...	n	
Behaviours	1	B _{1,1}	B _{1,2}	B _{1,3}	...	B _{1,j}	...	B _{1,n}	R ₁ = f(B ₁)
	2	B _{2,1}	B _{2,2}	B _{2,3}	...	B _{2,j}	...	B _{2,n}	R ₂ = f(B ₂)
	3	B _{3,1}	B _{3,2}	B _{3,3}	...	B _{3,j}	...	B _{3,n}	R ₃ = f(B ₃)

	i	B _{i,1}	B _{i,2}	B _{i,3}	...	B _{i,j}	...	B _{i,n}	R _i = f(B _i)

	M	B _{m,1}	B _{m,2}	B _{m,3}	...	B _{m,j}	...	B _{m,n}	R _m = f(B _m)
		M ₁ =f(B _{.,1})	M ₂ =f(B _{.,2})	M ₃ =f(B _{.,3})	...	M _j =f(B _{.,j})	...	M _n =f(B _{.,n})	
									M _R a=f(R _m)
									M _R b=f(M _n)
									M _R c=f(B _{m,n})

The forecasting method used in this thesis to forecast customer behaviour is based on internal behavioural data. In particular, these data are based on repeated observations from the same single act and from a group of customers. A detailed description of both the secondary data selected for the experiments and the forecasting process that has been carried out is provided in next chapters.

Summary

Forecasting, as one of the marketing research tasks which turns data into actionable information, is reviewed in this Chapter.

Although forecasting has been applied in several fields, the marketing area, and specifically food retailing forecasting interest has been noted. Despite the fact that there are several forecasting problems in retailing, there are many ways to solve these problems. Considering that, not only the most common forecasting methods have been reviewed, but also the types of data collection methods and tools (primary or secondary), the set of data used for the research (qualitative, quantitative, both) and the way of interpreting the data (logic or fuzzy logic) have been described as they are all likely to partially influence the final forecasting results. A general overview of the advanced forecasting techniques is also provided, in particular neural networks advantages and disadvantages as they are very similar to the fuzzy learning technique used in our experiments (See Chapter 4).

From all the possible retailing forecasting interests, forecasting customer behaviour is chosen as a key part of this research. Although there are several publications in the literature aiming to forecast customer behaviour, the majority of them are based on attitudinal data or intentions, but a few are supported by behavioural data. Then, and based on the customer research framework described in Section 2.7, the aim of this thesis is forecasting customer behaviour using behavioural data.

Nowadays, the majority of supermarkets use checkout laser-scanner technology to capture data at point of sale. There is the opportunity to use part of this collected

secondary data to forecast future behaviours. However, the problem is that the volumes of raw data generated in a superstore are enormous, and not always easy to deal with (Blundin, 1998). Sometimes, new methods are needed, which includes AI systems (Baron and Lock, 1995).

CHAPTER 4

Artificial Intelligence subfields: Machine Learning and Fuzzy Logic

Introduction

Artificial intelligence (AI) is a relatively new field studied and influenced by various domains. The fact that different scientists think of AI differently affects not only the way the area has evolved but also it has created certain confusion about the vast terminology employed. Based on that, a general overview of AI is provided in the following section. In particular, a general description of the history of AI is first explained, the reasons which explain the lack of an unified definition of AI are described and a definition selected. The history of AI in business is also presented in this section. Furthermore, a note on terminology is added. The machine learning and fuzzy logic basics are also introduced in this chapter as they provide the framework for the development of the empirical research of this thesis.

4.1 AI: A general overview

AI was born in The Dartmouth Conference, chaired by McCarthy in 1956 (Russell and Norvig, 1995). McCarthy wanted to bring together all the people he knew who had an interest in computer 'intelligence'. He knew there were several independent laboratories working in some areas of the field, but they had never shared their knowledge. The Dartmouth conference had an important effect. What had previously been a scattering of individual enthusiasts working in isolation was suddenly a scientific community, sharing not only their know-how, but also their processes and goals.

In the following years, artificial intelligence laboratories were established at a number of universities - Carnegie Mellon, in charge of Newell and Simon, at Stanford, under McCarthy, MIT under Marvin Minsky and at Edinburgh under Donald Michie.

When analysing the historic evolution of modern AI, three stages may be distinguished.

The first one, from Dartmouth to 1975, is the most symbolic period because computers would allow a change in the way researchers were simulating the human cognitive approaches. This is the stage where the knowledge connectionist representations started, when the first neural networks were designed and when symbolic programming languages were adapted to IA. The fuzzy sets theory was founded in 1965 by Zadeh (Cox, 1994; Klir and Yuan, 1995; Kosko, 1995; Dimitrov, 2002).

The second stage is from 1976 to the mid 1980s. During this stage the diverging specialities across the AI field emerged. It is considered as the period of the knowledge systems. Researchers commonly showed a considerable interest about understanding natural language, automatic problem solving and scenario analysis. At this stage also expert systems are introduced and analysed in depth (e.g. DENDRAL, MYCIN¹, PROSPECTOR) and the distinction between knowledge bases and inference engine is arisen (Russell and Norvig, 1995). These several research interests allowed a better understanding of heuristic search problems and knowledge representation (Brooks, 1991). Therefore, at this stage, there is an increasing trend to focus on processing knowledge instead of focusing simply on processing data (Russell and Norvig, 1995). In reference to commercial applications, the first industrial application of a fuzzy logic was introduced in a fuzzy predictive system to control cement manufacturer (Kosko, 1995). During the 1980s commercial AI startups were created by companies and academics.

During the 40 years after Dartmouth, the central AI studies were virtually focused on search methods (e.g. simulating human reasoning) and knowledge representation

¹ A revision and explanation of not only the most popular early expert systems such as MYCIN, DENDRAL, MACSYMA, PROSPECTOR but also the recent ones (XSEL, GENESIS) is developed by Harmon and King (1985). In addition, further explanation about expert systems is provided by Jackson (1990).

techniques. During this period of time it was demonstrated that intelligence not only needs the ability to reason but also information about the external world.

In the third stage, expert systems (ES) started to show their limits. AI scientists realised that an intelligent machine not only has to overcome search and knowledge representation tasks, but also it needs to examine its environment, decide the best actions, communicate, and learn about its mistakes (Alpayin, 2004). Therefore, vision, speech systems, planning, natural language or learning methods were also being studied to solve AI problems. It is important to note the World Wide Web emergence and the development of supercomputers such as Deep Blue (Korf, 1997) improved both the speed and the capacity to access any type of information (Bigus, 1996).

During the 1990s, there was a resurgence of interest in neural networks. There were several reasons for this, including the appearance of faster digital computers on which to simulate larger networks, the interest in building massively parallel computers and most importantly, the discovery of new neural networks architectures and powerful algorithms (Rich and Knight, 1991).

In recent years the development of AI has encompassed the stage of the intelligent agents. AI attempts to spread its research subfields through joining several methods applied in the intelligent systems design. These systems are generally characterised to model the information one agent knows from another agent. Consequently, they need to function depending on the common knowledge about the environment (Wooldridge, 2002).

Although these 50 years of investigations have been very productive, it is important to highlight AI is still a new and fast-developing area, and this constant evolution makes it difficult to define the concept of Artificial Intelligence, and consequently, specify its boundaries.

A detailed chronological explanation of the main events about the history of AI is

provided by Russell and Norvig (1995). From pages 16 to 25, the authors not only went back until the 5th century, demarcated as Ancient History of AI², but also propose future trends of the field. Also Newell (1993) explores the intellectual issues in the history of AI.

4.1.1 Defining AI

Despite the fact that there is a dictionary definition, there would appear to be no accepted unified meaning for Artificial Intelligence. The lack of a clear definition is due to 3 reasons.

Ongoing philosophical debate

Firstly, there is not an accepted academic definition of the term ‘intelligence’. This is a subject of much debate (Turing, 1950; Boden, 1977; Marr, 1977; Searle, 1980; Charniak and McDermott, 1985; Boden, 1990; Copeland, 1993; Simon, 1995). Most of the philosophers who considered the issues attempted to answer the question of whether or not a machine can think.

One way to answer this question is to apply the Turing Test (Turing, 1950). Turing’s proposal was to imagine you have a person able to communicate with two others, one male, one female simply by asking questions while those being questioned try to fool the interrogator about their gender. In the Turing’s test one of the human participants is replaced by a computer. If the computer can convince the interrogator it is human, Turing argued it can be said to be intelligent. On the contrary, Searle considers that computers will never be intelligent and he demonstrates his statement through the Chinese Room (Searle, 1984). He explains how the brain also has the mind, with feeling and emotions to influence thoughts. Moreover, he adds that a computer will never have feelings. He proves this argument with the Chinese room experiment, where he tries to show how a human running a computer program, may appear to understand but in reality they are just following instructions. Therefore, by following

² A chronology of significant events in the history of AI is found at <http://aaai.org/AIToics/hbhist.html#note>

instructions and producing correct outputs, a computer can simulate a human brain but will never duplicate it and its understanding. This debate about what is intelligence or not when referring to machines is explored in detail by Haugeland (1997).

According to Fischler and Firschein (1987), intelligence is easier to recognise than to define or measure. Intelligence is more an open collection of attributes than it is a single well-defined entity. Table 4.1 lists the attributes most closely identified with intelligence including learning, reasoning, understanding, linguistic competence, purposeful behaviour and effective interaction with the environment (including perception).

Table 4.1 Attributes of an intelligent agent

Attributes of an intelligent agent
Have mental attitudes (beliefs, desires and intentions)
Learn (ability to acquire new knowledge)
Solve problems (including the ability to break complex problems into simpler parts)
Understand, including the ability to make sense out of ambiguous or contradictory information
Plan and predict the consequences of contemplated actions, including the ability to compare and evaluate alternatives
Know the limits of its knowledge and abilities
Draw distinctions between situations despite similarities
Be original, synthesise new concepts and ideas, and acquire and employ analogies
Generalise (find common underlying pattern in superficially distinct situations)
Perceive and model the external world
Understand and use language and related symbolic tools

Source: Fischler, M.A. and Firschein, O. *Intelligence: The eye, the brain and the computer*. Reading, MA: Addison-Wesley, 1987:4

Moreover, there is a list of attributes (See Table 4.2) that despite the fact that they do not explain intelligence, they are considerably related. For instance, although emotions do not measure intelligence, being empathic and dealing with emotions is one type of human intelligence (Gardner, 2005).

Table 4.2 Human attributes related to intelligence

Human Attributes Related to, but Distinct from intelligence
Consciousness ³
Aesthetic appreciation
Emotion (anger, sorrow, pain, pleasure, love.)
Sensory acuteness
Motor skills

Source: Fischer, M.A. and Firschein, O. *Intelligence: The eye, the brain and the computer*. Reading, MA: Addison-Wesley, 1987:p4.

Multiple scopes

In general, the scientific community in the area shares the common accepted goal of AI which is to solve problems. Since the ability to solve problems is taken as a prior indicator that a system has intelligence, it is expected that much of the AI research is focused on building and understanding problem-solving systems (Newel and Simon, 1972). Nowadays, researchers are trying not only to investigate which are the suitable AI techniques to solve each type of AI problem, but also to attempt to solve problems traditionally not related to AI. The second reason, and related to the first reason is that there are several disciplines which influence and use aspects of AI to solve their research problems. Commonly, there is an intention to specify the boundaries of human intelligence, and finding the procedures to measure it. However, as the AI field extends from the intellectual area of engineers, computer scientists, psychologists, linguistics, mathematicians, biologists and philosophers, finding a suitable definition for all the disciplines is complicated. Based on that, a close relationship between AI and cognitive science, computer science and neuroscience is manifest in the literature. Considering the three main frameworks would help to pursue the research goal (Leake, 2002).

There is a connection between cognitive sciences such as philosophy or psychology and AI, documented in the literature (e.g. Sloman, 1978; Partridge and Wilks, 1990; Cummins and Pollock, 1991; Feigenbaum and Feldman, 1995; Haugeland, 1997).

³ Some scientists consider consciousness is an attribute of an intelligent agent.

'Cognitive Science is an interdisciplinary field that has arisen during the past decade at the intersection of a number of existing disciplines, including psychology, linguistics, computer science, philosophy, and physiology. The shared interest that has produced this coalition is understanding the nature of the mind' (AAAI, 2004).

It is difficult to determine the study of the mind when it has several meanings according to each discipline and to a specific period of time. However, the historical coincidence of several events is used to explain the link between cognitive science (CgS) and AI. Miller (1956) summarized numerous studies that showed that the capacity of human thinking is limited, with short-term memory. At this time, early generation computers had been around for only a few years, but the Dartmouth conference established the field of artificial intelligence, and promoted the development of programs that enable computers to display behavior that can (broadly) be characterized as intelligent. Then, behaviorism started to lose its influence in the 1950s (MIT, 2004a)

Philosophers and psychologists share a common goal which consists of specifying the boundaries of human intelligence and discovering the procedures to measure it. Consequently, they attempt to answer questions such as: What is a mind? (e.g. Anderson, 1983). What is consciousness? (e.g. Churchland, 1992). How humans acquire knowledge? (e.g. Boden, 1990). How do humans learn? (e.g. Fischler and Firschein, 1987; Hintzman, 1991; Hergerhahn and Olson, 1997). How does the memory and experience affect human behaviour? (e.g. Hintzman, 1991). What is intelligence? (e.g. Fischler and Firschein, 1987; Bernstein et al., 2003). Can a machine think? (e.g. Turing, 1950; Searle, 1984; Boden, 1990). Is it possible to measure intelligence? (e.g. Gardner, 2005) Are there different types of intelligence? (e.g. Gardner, 2005).

There is not a single answer to each question. Some authors are more optimistic when forecasting the machines' potential (McCarthy, 1990). Others are much more sceptical about the fact that a machine can 'own' human intelligence (Searle, 1984). However, both extreme opinions believe in the presence of an overlap between AI

and cognitive science (Berstein et al., 2003). Nowadays, cognitive scientists are focused on what happens in human brains during problem solving, remembering and perceiving (Berstein et al., 2003). However, AI not only is used to identify interesting and solvable information-processing problems, but also to solve them (Marr, 1977).

There is also a relationship between Computer Science (CS) and AI. Although there is not an agreement about whether AI is a branch of computer science (Leake, 2002) or not (Sloman, 2002), the bidirectional link to both fields is presented. Accordingly, AI uses CS and CS uses AI. Computer Science is the study of the phenomena surrounding computers (Newel and Simon, 1972). According to Sloman's standpoint, CS studies the forms of computation in general and their properties, both in hardware and in software. For instance, computer scientists study what programs can and cannot do, how programs should perform specific tasks efficiently and how programs use specific types of information. Under these conditions, the aim of building intelligence in computers is shared by both disciplines.

The existing relationship between neuroscience and AI is also evident. For instance, neural network models are inspired by neurons in the brain. Genetic algorithms simulate the process of natural selection of evolution. The retro feeding of both disciplines is mainly showed in section 4.2.

Strong versus weak AI

Thirdly, definitions over the literature seem to ignore the possibility of both strong AI and weak AI. As strong AI deals with the creation of some form of computer-based artificial intelligence that can truly reason and solve problems and be self-aware of that (Searle, 1984), weak AI deals with the creation of some form of computer based artificial intelligence that cannot truly reason and solve problems. Therefore, weak AI acts as if it was intelligent, but it would not possess sentience. According to Searle (1984) weak AI is defined as the view that the computer plays the same role in studying cognition as it does in any other disciplines. It is a useful device for simulating and studying mental processes, but the programmed computer

does not automatically guarantee the presence of mental states in the computer.

Discussing the validity, clarity or suitability of artificial and intelligence is not the aim of this thesis. As a result, and in order to avoid this possible philosophical debate, the following definition is chosen:

'Artificial Intelligence is the study of how to make computers do things which, at the moment, people do better' (Rich and Knight, 1991:13).

Although the definition has some limitations, it gives a broad view of AI and it avoids the philosophical issues that try to explain and define the actual meanings of artificial or intelligence.

4.1.2 History of Artificial Intelligence in business

In marketing, researchers and some companies also started to apply AI techniques. Although originally, the most common applied techniques were based on Expert Systems (ES), during the last 20 years some other AI subfields were applied as well.

'after almost 50 years of development, artificial intelligence products are ready to emerge from research lab and enter the real world of marketing research' (de Ville, 1997:38).

The main reason that explained the relative focus on expert systems was because they were easier to become commercialised, not because of their real utility. The high proliferation of ES applications led to useful and not so useful techniques. The failed applications generated a negative opinion, due to the lack of the expected utility. Moreover, it is important to note that as new subfields were applied and as more techniques based on AI were developed, ES became simpler to use.

There are relatively few applications of AI techniques in the marketing field. In fact, the majority of them are based on fuzzy logic systems and learning techniques, specifically neural networks (Venugopal and Baets, 1994; Fish, Barnes and Aiken,

1995; Vellido, Lisboa and Meehan, 1999; Desmet, 2001; Baesens et al., 2002; Boone and Rohem, 2002; Decker and Monien, 2003) and fuzzy neural networks (Kuo, Wu and Wang, 2002; Tseng and Tzeng, 2002).

Neural Networks were much researched during the 1990s, and peaked in 1996. For instance, Jurik (1993) provided a consumer guide to neural network software. Applications of neural networks in Management Science was also published by Krycha and Wagner (1999). During this decade, Tedesco suggested the term 'neural marketing', which is described as a flexible technique that has abilities to measure relationships of cause and effect and interpret expectations (Tedesco, 1992).

Interest in the analysis of user behaviour on the Internet has been increasing rapidly, especially since the advent of electronic commerce. The accessibility to huge updated datasets is possible thanks to the Internet. Therefore, experiments analysing the behavioural data to classify or identify patterns of electronic behaviour using machine learning techniques started to appear in the literature from the late 1990s. (Hendler, 2001; Paliouras et al., 2002).

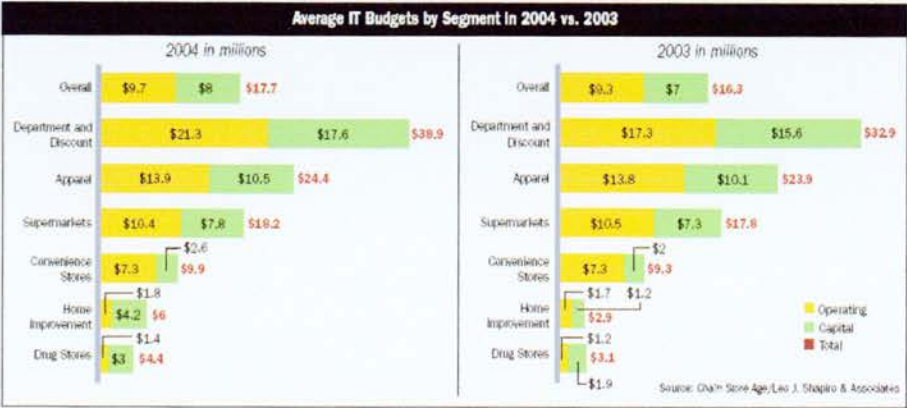
Although the interest in using different versions of Neural Networks is still increasing, there is also another technique which seems to be generating a similar interest among researchers which is genetic algorithms. Several papers focused on applying AI techniques in the marketing field introduce GA⁴ as the proposed methodology (Mitchell, 1996; Kim and Cho, 2000; Fish et al., 2004). During the 70s and 80s, expert systems were applied (Harmon and King, 1985). During the 1990s neural networks became more popular, and from 1997 on, the major researchers were moving to explore the use of genetic algorithms. Currently, the main trend is the combination of different AI subfields and AI techniques to solve virtually the same tasks.

Despite the fact that the average IT budgets spent by supermarkets in 2004 increased respect to 2003 (See Figure 4.1), and in particular, the area of business intelligence, forecasting and reporting which is not a retailer's priority also increased (Figure 4.2)

⁴ Genetic algorithms are briefly defined in section 4.2.3, Machine learning tasks and techniques.

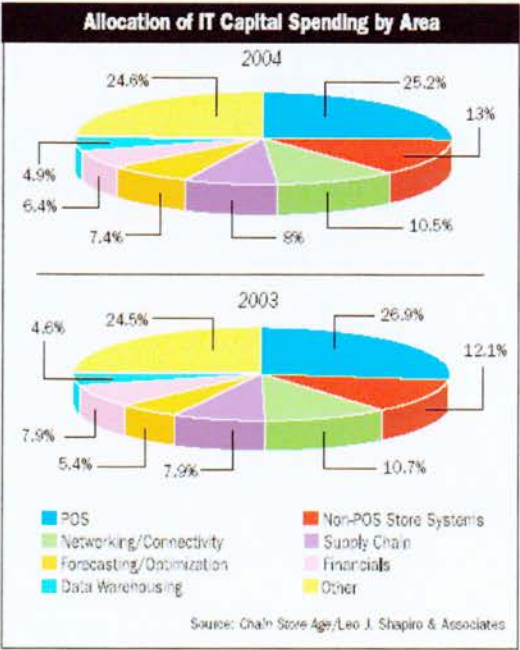
during the same period of time, it is difficult to determine the number of companies that are implementing AI methods in their business.

Figure 4.1 IT budgets by segment of the retail sector in USA



Source: Retail I.T. “23rd Annual Survey of Retail Information Technology.” *Chain Store Age*, 2004, November (2), 4A.

Figure 4.2 Capital spending on IT by retailers in USA



Source: Retail I.T. “23rd Annual Survey of Retail Information Technology.” *Chain Store Age*, 2004, November (2), 4A.

Moreover, there is no specific market research quantifying the percentage of companies that are using these innovative applications. Nevertheless, there is an unquestionable reality which is that AI has already penetrated the marketplace. For instance, as far as neural networks are concerned, it is known that although quite a few companies in diverse sectors have used NN, not so many use them in everyday business life (Borowsky, 1995).

Currently, some companies are considering the possibility of using AI methods to gain competitive advantage; others are just keeping an eye on AI developments. However, there are a noteworthy number of firms that are already implementing some AI techniques, particularly neural networks in some areas of management (Krycha and Wagner, 1999). A list of papers relating to the applications of NN in management is illustrated in table 4.3 (Vellido, Lisboa and Vaughan, 1999).

Table 4.3 References to NN applications in management

Management
<i>Competitive benchmarking</i>
Back et al. (1997)
<i>Decision</i>
Wilson (1994)
Kim and Park (1997)
Wang and Archer (1994)
Hill and Remus (1994)
Hill et al. (1994)
Tam (1994)
Willems and Brandts (1997)
Aiken (1997)
<i>Retail</i>
Thall (1992)
Belt (1993)

Source: Adapted from Vellido, A., Lisboa, P.J.G. and Vaughan, J. 'Neural networks in business: a survey of applications (1992-1998).' *Expert Systems with Applications*, 1999, 17 51-70.

Moreover, there are several companies which are focusing their core business in AI tools. For instance, Alyuda tools help business analysts to forecast, enhance the decision-making process, and discover knowledge from corporate datasets. Also

Clementine software from SPSS is used by retailers to analyse and predict information about their customers. Neural networks are being increasingly used for marketing applications, such as predicting customer demand and segmenting customers into well-defined categories (Neural Network Software, Zsolutions, Ward Systems, Neusciences⁵). According to Vellido, Lisboa and Vaughan (1999) data suppliers are usually private companies reticent to share any expert knowledge. This fact also explains the low number of publication in this topic.

4.1.3 A note on the terminology employed

The conventions observed and proposed by Whitby (1988:9-10) are followed in this Chapter:

'AI technology is used to cover all hardware and software used as part of any system which is described as AI'.

For instance: language programming, hardware devices, and algorithms and techniques are considered AI technology. Particular to technique description, a definition of AI technique proposed by (Rich and Knight, 1991:20) is also considered: An *AI technique* is the method that exploits knowledge (data).

'AI research is used for both research and development of AI technology and also into aspects of human psychology, such as the behaviour of system users, directly related to AI technology'.

Based on that, we would use *AI applications* to relate to this concept as well.

'AI based systems is used to denote computer systems which contain a large enough element of AI technology to make their behaviour significantly different from those which do not'.

⁵<http://www.alvuda.com/companyinfo.htm>; http://www.spss.com/training/contents/clem_intro.cfm; http://www.neuroxl.com/marketing_forecasting.htm; <http://www.tedescoanalytics.com>; <https://www.zsolutions.com>; <http://www.wardsystems.com/>; <http://www.neusciences.com/>; <http://www.fxpstock.com/>

Based on that, *AI solution* is used as an AI based system as well.

'Commercial AI is used to denote any AI-based system which is freely available for sale'.

Furthermore, when using *AI area/subfield*, it is going to refer the specific AI research carried out in the different main areas which AI is generally divided into. The classification is provided by Russell and Norvig (1995):

- Search,
- Knowledge capture, representation,
- Reasoning under uncertainty (fuzzy logic),
- Planning, vision and robotics,
- Natural Language processing,
- Machine Learning.

Other subfields/areas of research are also considered by AAAI which are:

- Genetic programming,
- Program synthesis,
- Artificial life,
- Artificial being,
- Distributed Artificial Intelligence,
- Swarm Intelligence.

Whatever the AI problem is investigated, all share the same characteristic which is complexity. Based on that, the majority of AI solutions cope with voluminous, changeable, fuzzy and dynamic data.

It is important to note that machine learning and fuzzy logic are the two AI subfields which are going to be analysed in this thesis as they delimit part of its framework. Consequently, learning theories, machine learning approaches, tasks and techniques are explained in section 4.2. Fuzzy logic fundamentals and applications are described

in section 4.3. Afterwards the synergies between machine learning and fuzzy logic proposed. Then, fuzzy learning applications in marketing are provided. Finally, the fuzzy learning algorithm (LAMDA) used for the empirical research is also described.

4.2 Machine learning

Prior to describing what ‘machine learning’ is and its major applications, a general overview of the term ‘learning’ is considered, as it is the most important part of the concept. Both the cognitive science and neuroscience (biological) perspectives about learning are described because they have been a clear inspiration to machine learning research.

4.2.1 Learning

‘Learning: is the alteration of behaviour as a result of individual experience. When an organism can perceive and change its behaviour, it is said to learn’ (Encyclopaedia Britannica, 2004).

Learning, as an attribute of human intelligence and a subfield of AI, can be explained and understood from several perspectives. However, it is necessary to note the psychological and neuroscience standpoints, as both have an influence to the fundamentals of machine learning.

Learning from a Cognitive Science perspective

From a psychological perspective, several approaches are found in the literature⁶.

As shown in Table 4.4, there are several basic ways to learn. For instance, people may learn from their genetically-endowed abilities to store knowledge, from supplied information or from past personal experiences. The discussion in section 2.1.2 is important at this point. As mentioned, most research in consumer behaviour has been oriented in investigating consumer choice (Mowen, 1995) which has been mainly

⁶ See Myers (2001) and Bernstein et al. (2003) for further explanation and details.

explained according two approaches which coincide with the major distinction in learning theory. Then, despite these previous six approaches in Bernstein et al. (2003), behavioural and cognitive learning perspectives are considered.

Table 4.4 Learning approaches to psychology

Approach	Characteristics
Biological	Emphasizes activity of the nervous system, especially of the brain; the action of hormones and other chemicals; and genetics. Therefore, behaviour is seen as the result of physical processes.
Evolutionary	Emphasizes the ways in which behaviour and mental processes are adaptive for survival. The most inherited aspects of behaviour are analyzed.
Psychodynamic	Emphasizes internal conflicts, mostly unconscious, which usually pit sexual or aggressive instincts against environmental obstacles to their expression when determining feelings and behaviours.
Behavioral	Emphasizes learning, especially each person-s experience with rewards and punishments when explaining behaviour.
Cognitive	A way of looking human behaviour that emphasizes mechanisms through which people receive, store, retrieve, process information and generates integrated patterns of action.
Humanistic	Emphasizes individual potential for growth and the role of unique perceptions in guiding behavior and mental processes

Source: Bernstein, D.A., Penner, L.A., Clarke-Stewart, A. and Roy, E.J. *Psychology*. Boston, USA: Houghton Mifflin Company, 2003: p20.

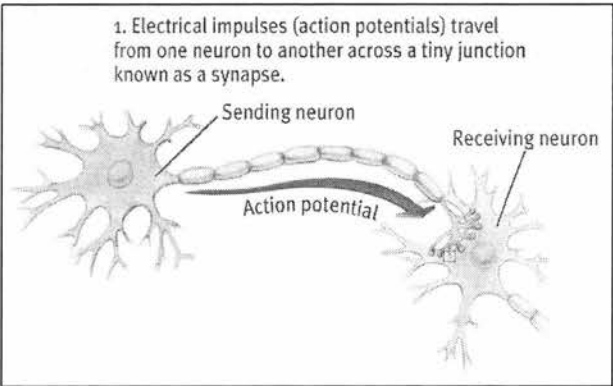
While behaviorism assumes that learning takes place as a result of response to external events, cognitive learning theories are based on the importance of internal mental processes that process information from the world to manage it. Both the behavioural and cognitive learning theories are based on the idea that people are born knowing very little and they absorb almost everything by the way of this general and ongoing process of learning (Charniak and McDermott, 1985; Fischler and Firschein, 1987; Winston, 1992).

Learning from a Neuroscience Perspective

The nervous system is a system of cells that allows an organism to gain information about what is going on inside and outside the body to respond appropriately

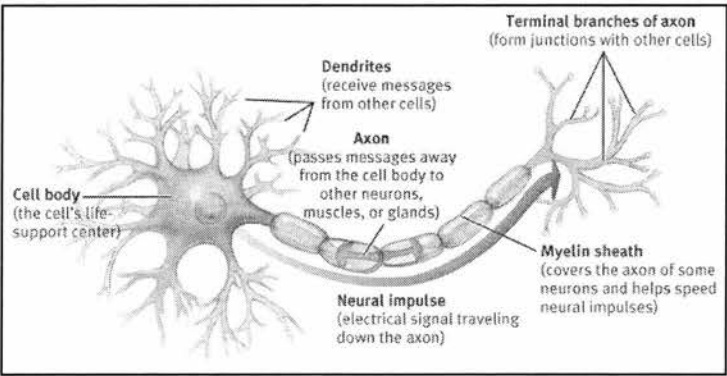
(Bernstein et al., 2003)⁷. The brain, as a major part of the nervous system, is composed of a large number of multi interconnected neurons. Figure 4.3 illustrates the synapse between two neurons. Synapse is the process which allows nerve cells to communicate with one another through axons (of one cell) connecting to dendrites (of the other), converting electrical signals into chemical ones.

Figure 4.3 Synapse between two neurons



Source: <http://www.wooster.edu/psychology/intro/synapse.gif> (march 2005)

Figure 4.4 The neuron



Source: <http://www.wooster.edu/psychology/intro/synapse.gif> (march 2005)

As illustrated in Figure 4.4, each neuron is a specialized cell which can transmit an electrochemical signal. The neuron has a branching input structure (the dendrites), a cell body (nucleus), and a branching output structure (the axon). When a neuron is

⁷ A deep explanation of how the nervous systems works is provided by Berstein et al. (2003) (Chapter 3).

activated, it fires an electrochemical signal along the axon. This signal crosses the synapses to other neurons, which may in turn fire. A neuron reacts only if the total signal received at the cell body from the dendrites exceeds a certain level (the firing threshold). The strength of the signal received by a neuron (and therefore its chances of firing) critically depends on the efficacy of the synapses (Berstein et al., 2003).

From a neuroscience perspective, learning consists principally in altering the strength of synaptic connections (Hebb, 1955). The mechanisms of learning and remembering seem to depend on relatively lasting variations and connections in the nervous system.

4.2.2 Defining machine learning

Based on the Cognitive Science (in particular cognitive and behavioural perspectives) and Neuroscience standpoint of learning, two considerations need to be taken into account. Firstly, note that machine learning does not aim to understand the processes underlying learning in humans or animals, but to build useful systems able to detect data patterns and learn about them. Watanabe (1985:1) defines a pattern

'as opposite of a chaos; it is an entity, vaguely defined, that could be given a name'.

For example, fingerprint image, human face, handwriting cursive words, or shopping behaviour could be various types of patterns. According to Winston (1992) they are two main kinds of learning. The first one is mixing new information to previously acquired knowledge. Usually, a great deal of reasoning is involved. This type of learning is related to analysing differences, managing multiple models, explaining experiences and correcting mistakes. The second type of learning is based on digging useful constancy out of variable data. Learning by recording cases, by building decision trees, by training different networks and by simulating evolution forms this group.

Machine learning believes there is a process underlying the generation of data,

although not all the details are known. Despite the fact that not always the process that explains the data we observe is going to be completely recognized, there is the possibility to detect certain patterns. According to Charniak and McDermott (1985), the essence of intelligence is the ability to learn,

'that is, to extract deeper and deeper insights into a repeated type of situation' (Charniak and McDermott, 1985:609).

The authors disagree with the previous mentioned western accepted learning approaches based on the idea that people are born knowing very little and absorb almost everything by the way of this general learn. Conversely, for an organism to learn anything, it must already know a lot (Charniak and McDermott, 1985). Learning begins with organized knowledge, which grows and becomes better organized. Based on this assumption, as much input data the machine has, most easy for the machine to learn and create knowledge.

Secondly, neither does machine learning aspire to study the complexity of the brain. The idea of building an intelligent machine out of artificial neurons has been around for quite some time. Some early results on brain like mechanisms were achieved by McCulloch and Pitts (1943), Hebb (1949), followed Rosenblatt (1958). Limitations of simple NN were slightly tested and NN architectures were improved (Rumelhart, Hinton and Williams, 1986). Several NN architectures appeared (Cruz, 1991). For more detailed discussion, comparisons between NN architectures and technical explanations are found in the literature (e.g. White, 1989; Hanson and Burr, 1990; Winston, 1992; Anderson, 1995; Anthony and Bartlett, 1999).

It is commonly accepted by the research community that there are two main cases where machine learning is necessary. Firstly, when human expertise does not exist or when humans are unable to explain their expertise. Secondly, when the problem to be solved changes in time, or depends on the particular environment. In this case, it is more suitable to design a machine learning system than designing a specific program for each particular circumstance.

Accordingly, based on the 2 previous considerations and taking into account the main cases when machine learning is necessary, the following definition is provided:

'Machine learning consists of programming computers to optimize a performance criterion using example data or past experience' (Alpaydin, 2004:3).

4.2.3 Machine learning tasks and techniques

During the 1990s, an explosion in machine learning was seen. According to Dietterich (1997), this explosion was due to the synergies obtained by the unification of several research communities that until that date had worked separately (for instance neural networks, statistics, pattern recognition, computational learning theory communities) and because the application of machine learning techniques to new types of research problems (including knowledge in databases, language processing and medical data analysis). Based on that, there is a wide range of learning tasks and learning techniques in the literature. To easily analyse them, Alpaydin's (2004) classification is followed:

- Learning Associations (e.g. association rules),
- Classification (e.g. discriminant, pattern recognition, supervised learning),
- Regression (prediction),
- Clustering (unsupervised learning),
- Reinforcement learning (learning actions).

In general terms, the goal of creating systems which can adapt to their environments and learn from their experience has, not only attracted researchers from AI, but also statisticians and mathematicians. Consequently, most of the mentioned tasks can be solved using statistical, mathematical or AI approaches. Therefore, it is normal that the same task or approach is termed in different words. For instance, supervised learning for continuous prediction is usually termed as regression and for nominal prediction is usually named as classification (Alpaydin, 2004). Therefore, prediction and classification are often called 'supervised learning tasks' because the training data include not only the input objects but also the corresponding output values.

On the other hand, to be intelligent, a system must learn and adapt to environmental changes. And not all the learning techniques over the literature are able to perform this way. Moreover, the degree of its sophistication is not the same. Therefore, the necessity to know what technique is most suitable for each problem also arises. According to Alpaydin (2004), machine learning techniques are classified according to these main groups:

- Decision Trees,
- Genetic Algorithms (GA),
- Induction and Case-Based Reasoning techniques (CBR),
- Neural Networks and Connectionists Systems,
- Pattern Recognition techniques.

Decision trees represent Boolean functions as a decision tree takes as input an object or situation described by a set of properties, and outputs a yes/no decision (Russell and Norvig, 1995).

Genetic algorithms are based on a biological inspiration. They are based on evolution concepts and include functions such as reproduction, mutation or competition.

'GA view learning as a competition among a population of evolving candidate problem solutions. A 'fitness' function evaluates each solution to decide whether it will contribute to the next generation of solutions. Then, through operations analogous to gene transfer in sexual reproduction, the algorithm creates a new population of candidate solutions' (Luger, 2002:471).

Case Based Reasoning is able to utilize the *specific* knowledge of previously experienced, concrete problem situations (cases). A new problem is solved by finding a similar past case, and reusing it in the new problem situation (Aamodt and Plaza, 1994).

Pattern recognition aims to classify data by means of a priori information extracted from a set of patterns. According to Watanabe (1985), given a pattern, its recognition

(classification) may consist of one of the following tasks, supervised classification which in the input pattern is identified as a member of a predefined class, or unsupervised classification in which the input patterns is classified into an unknown class. According to Jain, Murty and Flynn (1999), among the various frameworks in which supervised and unsupervised classification have been traditionally formulated, the statistical approach has been most intensively studied and used in practice. More recently, neural network techniques have been receiving increasing attention.

This thesis focuses on NN techniques and their ability to learn. Analysing each mentioned groups of techniques and their effectiveness in problem solving is not the centre of this work. Furthermore, historically both supervised learning (SL) and unsupervised learning (UL) have been the major tasks associated with NN. Therefore a description of just these major tasks is developed as well.

Neural Networks

Neural networks can be viewed as massively parallel computing systems with many interconnections (Anil, Duin and Mao, 2000). The basic model of a NN consists of a number of interconnected nodes. Each handles a designated sphere of knowledge, and has several inputs (which carry the values of variables of interest in the outside world) and one output (which form predictions, or control signals) to the network. Based on the inputs it gets, a node can ‘learn’ about the relationships between sets of data. The first NN was provided by McCulloch and Pitts (1943). Some years later, Rosenblatt (1958) proposed a very influential neural net model called the perceptron (Minsky and Papert, 1969). From the origin of perceptron networks to early 70’s several neurobiological breakthroughs were made (Cruz, 1991). The early 1980s were when the field entered a period of explosive growth. Several NN architectures are found in the literature during this NN explosion. The most popular ones include the Hopfield network (Hopfield, 1982), the back propagation network (Rumelhart, Hinton and Williams, 1986), the ART network (Grossberg, 1976; Carpenter and Grossberg, 1988) and Boltzmann Machines (Sejnowski and Rosenberg, 1986).

Neural network models attempt to use some organizational principles (such as learning, generalization, fault tolerance and distributed representation, and computation) in a network of weighted directed graphs in which the artificial nodes and their weights are connections between neuron outputs and neuron inputs (Anil, Duin and Mao, 2000). It is important to note that LAMDA also uses some of the organizational principles used by neural networks models such as learning, fault or manual tolerance and distributed representation and computation (See Section 4.6).

The main characteristics of neural networks are that they have the ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to and demonstrate to be useful in practical applications, along with the new trends and ideas (Alin, Duin and Mao, 2000). Based on that, NN are applicable in virtually every situation in which a relationship between the predictor variables (independents, inputs) and predicted variables (dependents, outputs) exists, even when that relationship is very complex and not easy to model in the usual terms of correlations or differences between groups. Moreover, one key feature of NN is that the relationship between inputs and outputs is learned by training. The training of the networks is performed by the learning algorithm (Burke and Ignizio, 1992). There is a considerable list of learning algorithms provided by the research community. The most popular is the backpropagation algorithm (Rumelhart, Hinton and Williams, 1986)⁸.

Supervised Learning (SL) and Unsupervised Learning (UL)

In SL the network is provided with training data. The training data consist of several examples of input data with the corresponding output. Based on these data, the network learns to infer the relationship between the input and output. Then, it is trained using some SL algorithm which uses the data to adjust the network's weights and thresholds. If the network is properly trained, it has then learned to model the (unknown) function that relates the input variables to the output variables, and can subsequently be used to make predictions where the output is not known.

⁸ A profound review of pattern recognition methods, approaches and classifiers is found in Anil, Duin and Mao (2000). Also bibliographic notes are recommended by the authors.

Unsupervised learning is performed when the system is given a collection of objects and it is asked to construct a model to group these objects. UL attempts to derive hidden structure from the raw data by discovering clusters in the input data, extracting features that characterise the input data and uncovering non-accidental coincidences within the input data (Wang et al., 2001). According to Everitt (1993), clustering is one natural type of unsupervised learning tasks as it aims to find and characterise clusters within the data. There are many introductory books on the topics of NN and connectionists systems. The bibliographical notes in Anthony and Bartlett's (1999) are highlighted as it collects an extended list of references. Particularly to UL, further explanations are provided by Barlow (1989) and Becker (1991).

4.3 Fuzzy Logic

Fuzzy logic is a concept introduced by Lofti Zadeh in 1965. The founder of fuzzy logic's proposal was focused on providing a logic system closer to the ways of inherent human thinking in order to deal with the complexity and uncertainty of the real life. A basic description and utilities of fuzzy logic is explained in this section.

4.3.1 Definition

Despite the fact that Zadeh is known as the father of fuzzy logic (Mansur, 1995), it is important to note that fuzzy logic resulted from several years of research. During the 1920s, Heisenberg demonstrated that even in physics, the truth of the statements is a matter of degree. Moreover, Lukasiewicz⁹ introduced the multivalued logic, which shows that there is a continuous option between the '0 or 1'. Russell used 'vagueness' when referring to multivalence logic. In 1937, Black published an article explaining the 'vague' sets. However, it was not until 1965 that Zadeh introduced the fuzzy set concept, based on multivalued logic and vague sets (Kosko, 1995).

⁹ Detailed explanation about multivalued logic is provided by Mansur (1995).

Until Zadeh's new logic, the majority of scientists followed Aristotle's logic, which is not only central in research but also in western language, education and way of living (Kosko, 1995). This is the A or not A statement. The bivalent reality of ones or zeros, the only two possible responses: 'True' or 'False'.

Scientists following this traditional logic attempt simplification. Therefore, the main objective is to find the lineal model which better matches with the real world, even if the world is not linear at all (Kosko, 1995). Although the statement which needs to be tested is not always exact and unchangeable, there is a motivation to turn it into being precise and static. According to Mandani (1977), mathematics as a whole was taken to be synonymous with precision. They believed that all the statements, when well defined, had to be whether true or false.

Before the 1970s, this bivalent logical approach was the main and common way to solve real problems. Virtually the majority of existing scientific theories did not have the capability to operate on perception-based information mainly because they based their meaning-representation on this predicated logic theory. In a departure from reliance on bivalent logical systems, an approach to commonsense reasoning based on fuzzy logic was developed for direct computing with human perceptions (Zadeh, 1992; 1996).

According to Dimitrov (2002:V),

'the nature of human perceptions manifests in their unique capacity to generalize, extract patterns and capture both the essence and the integrity of the events and phenomena in human life. This capacity goes together with an intrinsic imprecision of the perception – based information'.

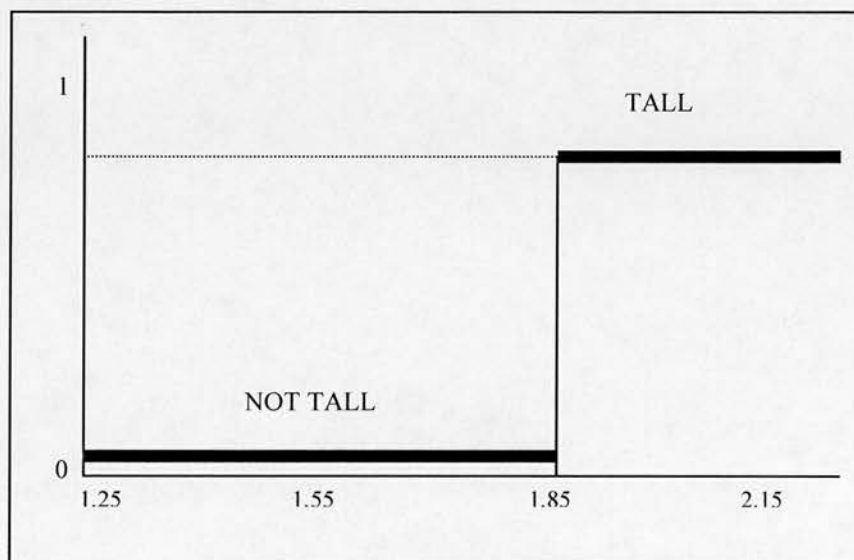
Based on this, fuzzy logical systems which could deal with information or knowledge that was imprecise, incomplete and not totally reliable were built.

4.3.2 Fuzzy set and fuzzy systems

The very basic notion of fuzzy systems is the presence of a fuzzy (sub) set. In classical mathematics there is a term, crisp set. A crisp set is any reunion of elements. By the characteristic function, every given element is assigned a number 1 or 0, depending on whether the element belongs to a particular subset or not.

On the other hand, a fuzzy set is a reunion of elements where the membership function can exhibit any real value between 0 and 1. In other words, a given individual can absolutely belong to a fuzzy set, absolutely not belong to it or show any in-between degree of membership (Zadeh, 1965). Zadeh explained fuzzy sets using the 'TALL man' example (Kosko, 1995). The traditional logic would illustrate the term 'TALL' according to the following graph (Figure 4.5). As it is illustrated in the figure, 'TALL' is decided to be at 1,85 metres. Less than this height the man is not tall. More than this height, the man is tall. The bivalence logic is evident in this example. The TALL and NOT TALL are two crisp sets.

Figure 4.5 Representation of crisp sets: TALL or NOT TALL

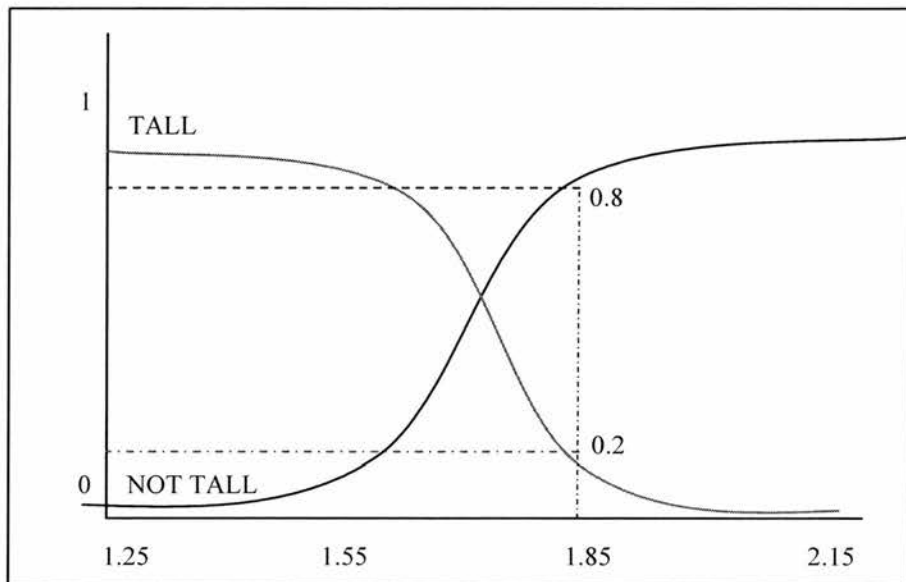


Source: Adapted from Kosko, B. *Pensamiento Borroso. La nueva ciencia de la lógica borrosa*. Barcelona, Spain: Crítica, (Grijaldo Mondadori, S.A.), 1995.

However, in the real world, people are not so strict when taking decisions. The world

is not always black or white but a range of greys. When referring to the range of greys, Zadeh used membership function. The membership function is a continuous function which shows the ‘membership’ of each specific height at the set TALL man. All men are TALL with in a certain degree. And all man are NOT TALL with in a certain degree. For instance, when analysing the TALL set from a fuzzy standpoint, the specific 1.85 m high belongs 0.2 to NOT TALL and 0.8 to TALL. In fact, the two functions coincide when A is equal to no A (See Figure 4.6).

Figure 4.6 Representation of fuzzy set: TALL and NOT TALL



Source: Adapted from Kosko, B. *Pensamiento Borroso. La nueva ciencia de la lógica borrosa*. Barcelona, Spain: Crítica, (Grijaldo Mondadori, S.A.), 1995.

The existing range of greys becomes totally evident when analysing customers’ patterns of behaviour. Terms such as frequency, loyalty, satisfaction, profitability, defections, convenience, price sensitivity, and redemption are just a few examples of ambiguous data exploited when analysing customers. Contrary to the information provided by the variable ‘sex’, which just has 2 answers ‘male’ or ‘female’, these mentioned examples have several interpretations. Similarly to the tall example, these ambiguous concepts can be represented in a fuzzy standpoint. Therefore, defining each concept according to fuzzy sets will help have a more realistic description of them. As Zadeh demonstrated, words announce set of elements. These elements

belong to each fuzzy set in a certain way. However, these are just sets. Sentences are groups of words. A sentence relates together groups of words, which is the basis of the human reasoning. A group of sentences is a fuzzy system (Kosko, 1995).

Defining the same operations as in classical sets, such as intersection and union is also possible with fuzzy sets and fuzzy systems. Zadeh (1965) suggested the minimum operator for the intersection and the maximum operator for the union of two fuzzy sets. Moreover, additional operations like inclusion, complement, relation and convexity are extended to such sets as well (Zadeh, 1965).

4.3.3 Characteristics of information

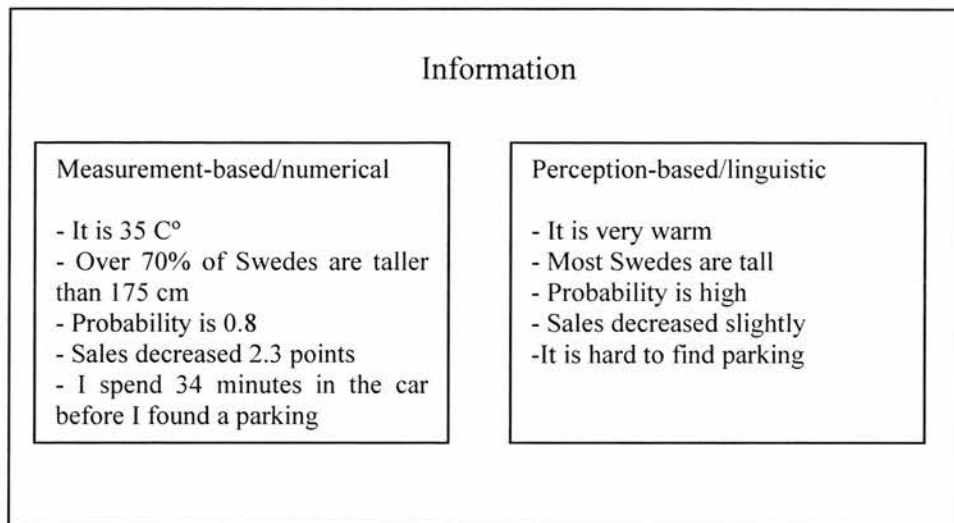
The human mind is a source of fuzziness. The human rational mind is not used to divide the whole reality into pieces in order to analyse, classify and label them. Once the parts of the reality are labelled, they are then put together to build a world that is not likely to be similar to the reality (Dimitrov, 2002). Information is imprecise and uncertain. Sometimes, it is incomplete as well. These possible features inherent to data are explained in detail.

The imprecise nature of the Information

There are two main types of information: numerical and linguistic types (See Figure 4.7). A linguistic variable is defined

'as a variable whose values are sentences in a natural or artificial language' (Zadeh, 1973:28).

Figure 4.7 Examples of numerical and linguistic information



Source: Adapted from Zadeh, L.A. "Fuzzy logic = computing with words." *IEEE Transactions on Fuzzy Systems*, 1996, 4 (2), 103-111.

Despite the fact that sometimes there is the possibility to measure the information, not always do humans have the tools to be precise about it. According to Zadeh (1996), numerical information may be viewed as a special case of perception-based information. Moreover, linguistic information is intrinsically imprecise, it is fuzzy.

The uncertainty of the information

The tradition in science to relate uncertainty with probability was broken with the Generalized Theory of Uncertainty (GTU), introduced by Zadeh (1973), as providing a broader perspective of uncertainty. Zadeh (1973) considers that there are 4 ways which underlie the way the attributes are related to the linguistic variables.

The first is the bounded ability of sensory organs, and ultimately the brain, to resolve detail and store information. For example, looking at customers, it might be seen that she is young (18 – 25 years) but her age as a single number cannot be identified. Secondly, numerical information is not likely to be available. For example, I may not know exactly how many customers have been lost to a firm, but my perception may be 'not many.' Third, when an attribute is not quantifiable. For example, degrees of customers' loyalty as: low, not high, high, very high, etc. are described because a

numerical scale is not applied. And fourthly, when there is a tolerance for imprecision which can be exploited through communication. For example, it may be sufficient to know that a customer is young; her exact age may be unimportant.

The information may be incomplete

The ability to collect, summarise and decide the relevant but incomplete information for solving a specific problem is an important asset of the human mind,

'as well as a fundamental characteristic that distinguishes human intelligence from the type of machine that is embodied in present-day digital computers' (Zadeh, 1973:29).

Complete lack of information will not support any decision making using any form of logic. But it is important to note that fuzzy logic allows decision making with estimated values under incomplete information (Munakata and Jani, 1994). For difficult problems, conventional nonfuzzy methods are usually expensive and depend on mathematical approximations, which may lead to a poor performance of the reality (Munakata and Jani, 1994).

Having explained the inherent features of information, the importance of disposing systems that are able to work with fuzzy variables are highlighted. Most of the basic tasks performed by humans do not require a high degree of precision. The human brain is able to discern between relevant and not relevant information. Based on that, the stream of information captured via the visual, auditory, tactile and other sense is reduced to the trickle that is needed to perform a concrete task. For instance, observing the shopping behaviour of a customer during a period of time, a manager could be able to know whether this customer would behave in a specific way according to a particular situation. Customers do not always behave in the same way. Customers are likely to change their 'habitual' behaviour according to the situation. Then, when analysing or predicting customers' behaviour automatically, an interval of flexibility is needed to cope with the ambivalence of behaviour, a fuzzy perspective it is also suggested.

Fuzzy logic provides approximate but consistent solutions to complex problems, where numerical data usually are noisy and incomplete and the linguistic information is imprecise and vague (Tchamova and Semerdjiev, 2002). According to Zadeh (1973), digital computers failed when trying to deal with the reality of human thinking and behaviour. Consequently, he proposed an approach which shifted from the precision, rigor and mathematical formalisms towards the partial truth. Even Zadeh proposes equivalence between fuzzy logic and computing with words, which is a methodology in which words are used in place of numbers for computing and reasoning (Zadeh, 1996; Kosko, 1995; Yakoubov and Haberman, 1998; Dimitrov, 2002).

In general, the employment of fuzzy logic might be helpful (Smithson, 1987), for very complex processes when there is no simple mathematical model, for highly nonlinear processes or if the processing of (linguistically formulated) expert knowledge is to be performed.

4.4 Fuzzy machine learning hybrid systems

Fuzzy logic is 40 years old and was considered one of the main AI subfields until the 1990s. An important part of fuzzy logic research was focused on approximate reasoning and reasoning under uncertainty (Dubois and Prade, 1995).

Once the research community started to realize the potentiality of fuzzy logic, several researchers, applying combinations of approaches, started to be published in different fields. As Flach (2001) states, when analyzing the state of the art in machine learning, there is a clear trend to research which combines approaches that previously were considered as separate approaches. It is noteworthy that two areas which became conscious of their potential synergies were reasoning and learning. According to Kosko (1995), the potentiality of fuzzy learning systems is not only the capability to reason as humans but also to learn from the experience with imprecise, uncertain and incomplete information.

Across the literature, a conclusion can be drawn that there is an increasing trend to combining forms of machine learning and fuzzy logic. For instance, the number of publications that introduce fuzzy learning techniques from year 2000 until 2005 is considerably higher than the publications between 1995 and 2000. In particular, according to Munakata and Jani (1994), the most active trend is various forms of hybrid systems of fuzzy logic and other areas such as neural networks and genetic algorithms. Several applications have shown the advantages of developing fuzzy machine learning techniques in many domains. Despite the fact that most of these applications started in fields such as medicine, biology, ecology and informatics, some applications are also found in marketing.

4.5 Fuzzy learning applications in marketing

Despite the fact that there are few applications of AI in marketing compared to fields in mainstream science, it is interesting to notice that the majority of AI applications over the marketing literature are based on machine learning and fuzzy logic approaches. There are few publications found over the literature. Moreover, not all the research works with the same technique. Therefore, a classification across different marketing topics is provided in Table 4.5 which includes:

- market response prediction and strategy development,
- market segmentation and targeting,
- sales forecasting,
- choice prediction and customer interest,
- and basket analysis.

Moreover, Table 4.5 also shows whether the research work used a learning technique (most of them use NN) or a fuzzy learning technique.

Table 4.5 Learning and fuzzy learning applications in marketing: Literature review

	Machine learning (NN)	Fuzzy learning technique
Market response prediction and Strategy development		
Baets, Brunenberg and Van Wezel (1998)	v	
Li (2000)		v
Duan and Burrell (1995)		v
Market segmentation and targeting		
Mazanec (1992)	v	
Dasgupta, Dispensa and Ghose (1997)	v	
Fish, Barnes and Aiken (1995)	v	
Davies, Moutinho and Curry (1996)	v	
Balakrishnan et al. (1996)	v	
Setiono, Thong and Yap (1998)	v	
Ha and Park (1998)	v	
New product acceptance	v	
Kumar, Rao and Soni (1995)	v	
Huang and Lippman (1997)	v	
Tam and Kiang (1992)	v	
Yoon, Swales and Margavio (1993)	v	
Hruschka (1986)		v
Wedel and Steenkamp (1991)		v
Ozer (2001)		v
SteenKamp and Wedel (1992)		v
Sales forecasting		
Hruscka (1993)		v
Dutta, Shekhar and Wong (1994)	v	
Shanker, Hu and Hung (1996)	v	
Kuo, Wu and Wang (2002)		v
Tseng and Tseng (2002)		v
Heshmaty and Kandel (1985)		v
Choice prediction and customer interests		
West, Brockett and Golden (1997)	v	
Baesens et al. (2002)		v
Hamuro et al. (2001)		v
Araya, Silva and Weber (2004)		v
Wray, Palmer and Bejou (1994)		v
Basket Analysis		
Decker and Monien (2003)		v
Desmet (2001)		v

Market response prediction and strategy development

There is a research interest oriented to designing models which may be used as a support decision tool for managers (Dimitrov and Wechler, 1975). According to some authors, the combination of a neural network approach and fuzzy logic when interpreting data implies a considerable improvement in the utility of these models (e.g. Li, 2000; Duan and Burrell, 1995).

Market segmentation and targeting

Targeting is one of the main interests of the marketing community (Dibb and Simkim, 1991). Not only is there research attempting to demonstrate whether neural networks are more accurate than statistical methods (in particular to discriminant analysis) when targeting (e.g. Huang and Lippman, 1987; Tam and Kiang, 1992; Yoon, Swales and Margavio, 1993; Fish, Barnes and Aiken, 1995), but also there is research that highlights the efficiency of fuzzy learning techniques for solving the same tasks (e.g. Hruschka, 1986; Wedel and Steenkamp, 1991; Ozer, 2001; Steenkamp and Wedel, 1992). Market segmentation resulting from the application of NN techniques are tested by several researchers (Mazanec 1992; Dasgupta, Dispensa and Ghose, 1997; Davies, Moutinho and Curry, 1996; Balakrishnan et al., 1996; Setiono, Thong and Yap, 1998).

Sales forecasting

Not only NN have been applied to sales forecasting (e.g. Dutta, Shekhar and Wong, 1994; Shanker, Hu and Hung, 1996) but also fuzzy neural networks have been applied as well (Kuo, Wu and Wang, 2002). Moreover, fuzzy regression techniques are tested as well in this area (e.g. Tseng et Tseng, 2002; Heshmaty and Kandel, 1985).

Choice prediction and customer interests

Retail firms analyse their past sales data to learn their customers' behaviour in order to improve customer relationship management (Flach, 2001). Firms share a common goal which is to identify prospective customers. Despite the presence in the research literature of several traditional methods to classify customer interests, new techniques have been applied in this area which mix learning and reasoning (e.g. Venugopal and Baets, 1994; Wray, Palmer and Bejou, 1994; Baesens et al., 2002; Hamuro et al., 2002; Araya, Silva and Weber, 2004).

Market basket analysis

Market Basket analysis attempts to find patterns of products which may help

managers taking appropriate decisions. Decker and Monien (2003) and Desmet (2001) provide learning approaches to solve this problem. However, these studies do not note the implicit fuzzy approach.

It is important to note that most of these researches have a common denominator which is the capability of accessing secondary data. The majority of these publications use the data from scanner systems, e-commerce track or external secondary data. This is a key point because the greater the amount of historical data, the more useful is the applicability of AI approaches (Charniak and McDermott, 1985), and the more successful the results are likely to be. In general, most of the authors mentioned in this section suggested their particular technique is still under development when solving the specific research problem. In the same way, a fuzzy learning technique is particularly used for our empirical research. The experiments developed in this thesis are basically related to forecasting customers interests and targeting.

4.6 LAMDA, a fuzzy learning algorithm

LAMDA (Aguilar-Martin and Piera, 1986; Piera and Aguilar-Martin, 1991; Aguado, 1998) is a classification method based on hybrid connectives (Aguilar-Martin and Piera, 1986). These hybrid connectives allow the several forms of partial information, resulting from the relationship between the individual to each variable (marginal adequacy degree) to be turned into a simple result which assigns each individual to an appropriate existing segment. Then, there are as many marginal adequacy degrees as variables (predictors) used in the experiment. Moreover, there are as many global adequacy degrees (GAD) as the number of existing segments considered in the experiment. *LAMDA* gives the possibility of forecasting based on supervised learning and unsupervised learning.

4.6.1 Theoretical aspects of LAMDA

As it has been said, in *LAMDA* methodology, it is assumed that a Marginal

Adequacy Degree ($MAD(X_i/P_i)$) can be evaluated for each descriptor, and consequently a global adequacy degree ($GAD(X/P)$) is obtained by a logical combination, or connection operator L , such that

$$GAD(X/P) = L [MAD(X_1/P_1), MAD(X_2/P_2), \dots, MAD(X_n/P_n)]$$

The operator L is defined as a linearly compensated mixed connective (Piera and Aguilar-Martin, 1991), for instance the interpolation between a t-norm and a t-conorm,

$$L = (1-\beta) T + \beta \perp$$

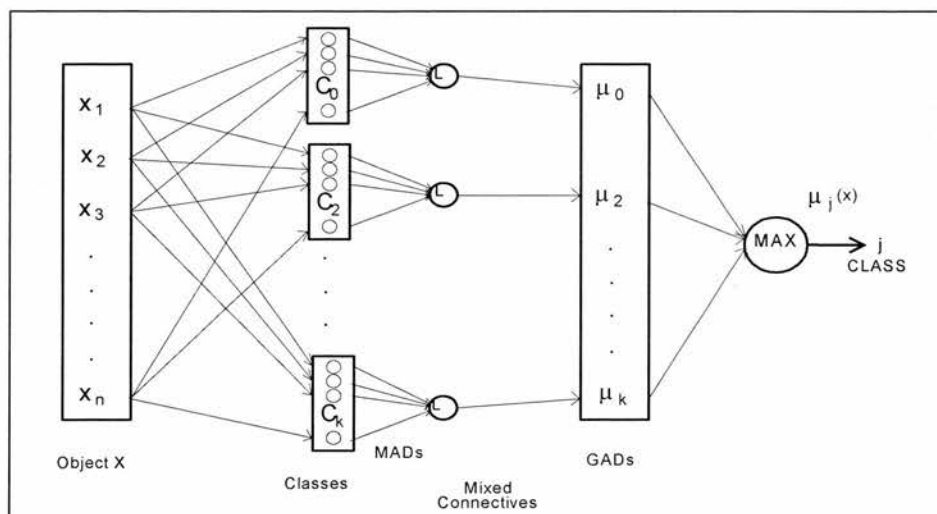
by means of the β parameter such that $\beta = 0$ represents the t-norm, T , for example the Minimum and $\beta = 1$ means the t-conorm, \perp , for example the Maximum. This parameter will -inversely- determine the exigency level of the classification, so it can be called tolerance. There are multiple examples of t-norms and their t-conorms. However, LAMDA algorithm just uses 4 which include MinMax, Probabilistic, Frank and Lukasiewicz. These operators are termed as connectives.

To clearly show how connectives work, we can think that by applying a t-norm we are employing the maximum exigency, and an individual (customer) will be considered only as a member of a class (pattern of behaviour) if every variable indicates so. If we opt for the t-conorm, the minimum exigency or maximum tolerance is chosen, and a customer will be considered a member of a pattern of behaviour only if one descriptor or studied variable indicates so. Therefore, by adding the parameter β , that determines inversely the demanding degree of the classification, we can determine the necessary tolerance for each particular case. A hybrid connective takes places from the interpolation resulted between a t-norm and its t-conorm, according to the parameter β .

Figure 4.8 illustrates how LAMDA algorithm works. From the marginal adequacy degrees, the GADs are calculated by means of the hybrid connective. In general, each individual is assigned to the segment which presents the highest GAD.

However, the GAD to each segment is also known.

Figure 4.8 LAMDA architecture



Source: Aguilar, J., Piera, N. and Aguado, J.C. "Data Based Inductive Knowledge Representation." In *International Panel Conference on Soft & Intelligent Computing*, Budapest, October (1996):p4.

It should be noted that this classifying structure bears a remarkable resemblance to an artificial neural network. Thus, not surprisingly, there is a training stage when classes can be modified and created and a pure pattern recognition step when all that is needed is to assign individuals to fixed classes. The result is a fuzzy partition of the data set. Nevertheless, these stages can be compatible and learning could go on forever. To guarantee that no classification decision is made with insufficient membership, a 'Non Informative Class' (NIC) is introduced; this NIC class must accept any threshold of adequacy. The choice of this threshold may be either imposed by the marginal adequacy functions chosen, or introduced by and expert.

Whenever the object to be classified does not possess membership in another class that is greater than its membership in the non informative class (NIC) it can not be properly assigned because it is not close enough to any existing class. When performing recognition from supervised learning, this means either that the individual will not be classified or that the resulting classification will not be relevant. But if the algorithm is learning without supervision, this situation implies

that a new class is needed, and it will be automatically constructed between the new individual and the NIC. In any case, the NIC always remains empty.

To start working with LAMDA program first, it is necessary to define a context. A context is defined when the selected variables to solve the research problem are identified. LAMDA needs to know how many descriptors are going to be used. Moreover, LAMDA algorithm can deal simultaneously with qualitative and quantitative data. Therefore, the number and type of predictors need to be known as well. Before LAMDA's learning tasks is started, the range of the quantitative variables and the categories of the qualitative variables have to been introduced to the programme as well. LAMDA's context may be built in two different ways. Firstly, it can be induced from the input data. The user may also introduce the context directly.

As mentioned, LAMDA is able to perform supervised learning and unsupervised learning. In both cases, the hybrid connective must be selected and the tolerance as well. There is the possibility to choose a degree of tolerance either automatically or manually.

In addition, the number of iterations of the learning process has to be also decided. When a maximum number of learning iterations is established, the number of times that LAMDA repeats the analysis of the individuals is prefixed. The segments resulting from the learning process tend to look for stability, which means that the individuals located in each segment tend to remain in it. The prefixed number of iterations avoids the fact that the learning and recognition process lasts an excessive time compared to the quality of the results.

In the supervised learning, as the final goal is always to mimic the discrimination capabilities of an expert, the quality of the results can be tested immediately by comparing LAMDA results with the 'right' classification provided by the expert. The supervised learning task of LAMDA is experimented in Chapter 5.

Differently to the supervised learning, unsupervised learning can not be tested and has to be employed to make new suggestions to the expert. LAMDA can automatically generate segmentations induced from the available data, often producing segmentation systems the expert never thought of (some of which may prove very useful in a dynamic and creative field like marketing). To choose which is the most suitable classification for the forecasting, a criteria needs to be predefined by the expert. A detailed explanation of the unsupervised learning task is found in Chapter 6.

Summary

A wide range of learning tasks and learning techniques are found in the literature. According to Dietterich (1997) this explosion appeared during the 1990's mainly due to the unification of several research communities and the application of machine learning techniques to new types of problems and fields. From all the machine learning techniques, NN are described in this chapter. In addition, from the entire machine learning tasks, supervised and unsupervised learning has been highlighted as well.

Analysing learning from a neuroscience perspective helps to explain how neural networks work, as they attempt to imitate the neurons connections in the brain. To better understand the learning tasks, the learning theories from a Cognitive Science perspective are noted. Despite the fact that consumer choice has been mainly studied from a cognitive perspective (Mowen, 1995), a behaviourism approach is followed in this work. Behaviourism theories assume that learning takes place as a result of response to external events. Based on this, and related to the previous section 2.6.1, when related to groceries, customer and shopping behaviour are likely to be better understood by the behavioural approach as performing a specific behaviour problem solving is virtually excluded.

Furthermore, it has been highlighted that customers' patterns of behaviour are the focus of learning. Customers are not simple robots. People tend to behave differently

according the situation so their patterns of behaviour are not always certain and rigid. To deal with this dynamic, uncertain and complex aspect, fuzzy logic is considered as it is based on the concept of partial truth.

Finally, LAMDA is described for the experiment. LAMDA's architecture is similar to the basic model of neural networks. Moreover, it is based on fuzzy sets theory. Then, its capabilities for learning and dealing with unambiguous data are analysed and considered to be suitable to forecast customer behaviour. The following chapters describe the empirical research applying LAMDA's methodology.

CHAPTER 5

Experiment 1: Forecasting customer behaviour in the Spanish grocery industry: Identifying the customers who are going to defect.

5.1 General research lines

Within the research literature, forecasting customer defection has been studied from three directions. It is important to note that not only are there no rigid boundaries between the three lines but also they are complementary. Nonetheless, as the existing research literature is very extensive, distinguishing between these three main lines is useful as a way to provide a framework for the background to the first experiment.

The first line considers forecasting defection as a type of information required to formulate an appropriate strategy for customer retention. Despite the fact that information about forecasting defections is relatively attractive for this line of research, the major interest for academics and practitioners within this perspective is to finding out whether companies that focus their efforts on retaining customers are more profitable than the ones that attempt to capture new ones (e.g. Dawkins and Reichheld, 1990; Reichheld and Sasser, 1990; Rust and Zahorik, 1993; Storbacka, Strandvik and Grönroos, 1994; Reichheld, 1996b; Mozer et al., 2000; Jones, Mothersbaugh and Beatty, 2000). Researchers in this area are also concerned with providing customer retention models (e.g. Rosenberg and Czepiel, 1984; DeSouza, 1992; Ahmad and Buttle, 2002; Tao and Yeh, 2003) as a useful tool for managers.

The second line of research attempts to understand the reasons why customers defect, measuring the influence of the different reasons. There are several studies focused on explaining the factors that influence defections. Traditionally, customer satisfaction has been the predominant metric used by firms for detecting customer defections (e.g. Rust and Zahorik, 1993; Mittal and Lassar, 1998; Mozer et al., 2000). The measurement of customer satisfaction has been linked to repurchase intentions (e.g. Coyne, 1989; Reichheld and Sasser, 1990; Anderson and Sullivan, 1993), and

repurchase behaviour (e.g. LaBarbera and Mazursky, 1983; Bolton, 1998). Although customer satisfaction is used to explain why customers defect, it also has some limitations. For instance, high satisfaction does not guarantee that customers do not defect (Reichheld, 1996a; Capraro, Broniarczyk and Srivastava, 2003). There are other factors that may explain and influence customer's defections such as the notion of the psychological state of loyalty (Oliver, 1999), switching barriers (Day and Bodur, 1978; Gwinner, Dwayne and Bitner, 1998; Capraro, Broniarczyk and Srivastava, 2003; Ranaweera and Prabhu, 2003), customers' level of knowledge about the available alternatives (Capraro, Broniarczyk and Srivastava, 2003), frequency of purchasing (Schmittlein and Peterson, 1994), length of relationship between customer and company (Anderson and Weitz, 1989; Verhoef, Franses and Hoekstra, 2002), customer demographics (East et al., 1995; McGoldrick and Andre, 1997; Buckinx et al., 2004), complaining customer voices (Roos, 1999) and industry characteristics (Roos, Edvardsoon and Gustafsson, 2004). All have been suggested as factors to explain defection.

Apart from analysing the potentiality of each factor when explaining the likelihood of defections, it is important to note that some authors attempted not only to identify these reasons (e.g. Keaveney, 1995) but also to classify them. Roos (1999) provides the critical switching path model (CSP) which consists of a switching process that starts with the customer's awareness of some negative aspects in the relationship with his/her supermarket and finishes in some cases in a irrevocable switching decision.

When analysing the CSP, a sequence of various combinations of key factors appears. Roos, Edvardsoon and Gustafsson (2004) mainly distinguish between the underlying factors (triggers) and the expressed factors (switching determinants).

The trigger indicates the sensitive factors influencing customer behaviour change which represent the reasons why customers begin to consider switching. The authors mainly distinguish between situational, influential and reactional triggers.

- Situational triggers are defined as changes in the customers' own life.
- Influential triggers are defined as competitor's effort to increase their market share.
- The reactional trigger comprises the critical incidents in interactions between customers and service providers.

Switching determinants are also grouped in 3 categories, pushing determinant, swayer determinant and pulling determinant.

'Pushing determinant is the switching determinant that is perceived by the customer as the reason for switching to another supermarket (price, range of foods, location). Swayer is the determinant which can either mitigate or prolong the CSP (personnel, price, range of goods, location, habit, queuing, variation, design, atmosphere, policy. Pulling determinant explains why the customer returns after switching' (Roos, 1999: 74-75).

Hence, it seems that a classification of the factors depending on whether they are particular to customers, resulted from the relationship of customer to store or whether they are special features from the industry, would be appropriate.

The grocery retail environment is characterised by intense competition that allows customers to have a wide array of alternatives.

'Perceived switching barrier is defined as the consumer's assessment of the resources and opportunities needed to perform the switching act, or alternatively, the constraints that prevent the switching act' (Ranaweera and Prabhu, 2003:379).

According to the definition, it is also demonstrated that the higher level of perceived switching barriers (and consequently perceived switching costs), the less the number of customers defections.

Within the retailing environment, in general, customers do not perceive high switching costs when changing their supplier. As grocery customers have a non-

contractual setting with the company, they have the opportunity to change their purchase behaviour without informing the company about it (Buckinx and Van den Poel, 2005). This may be one of the reasons to explain why they split their purchases in several competitive companies (Buckinx and Van den Poel, 2005).

The third line of research attempts to predict defections, based on customer analysis. To predict defections, different forecasting methods are used. Basically, they can be split between quantitative and qualitative methods. The main distinction between them is the type of data used for the forecast. The first type of forecasting methods is usually based on more behavioural data whilst the second type is mainly based on purchasing intentions.

Although there are some churn analyses in the literature, it is important to note that defections in the food retailing sector have been under researched (Buckinx and Van den Poel, 2005), particularly attempts to forecast defections.

As Buckinx and Van den Poel (2005) stated, not only is it difficult to find research based on retailing defections but also it is very difficult to find partial defection analysis, whatever the sector. In fact, the majority of the studies are focused on predicting total defections. Total defections are easier to predict, particularly in those sectors where there is a contractual setting between the customer and the company.

5.1.2 Total versus partial defections

Whatever the main research interest, it is important to note that directly or indirectly, most of the publications within the three perspectives tend to distinguish between total and partial defections.

Total defection occurs when customers totally interrupt their relationship with the company (Buckinx and Van den Poel, 2005).

Partial defection takes place when buyers do not defect from the company suddenly. They switch some of their purchases to another store (Buckinx and Van den Poel,

2005). According to Bloemer et al. (2003:118) partial defectors are the customers who present a high probability to defect in the near future as

'they are the customers who, although are loyal at the moment, might have a higher propensity to leave than others and are somehow less willing to stay in the relationship'.

According to Roos, Edvardsson and Gustafsson (2004), the outcome stage comprises the kind of switching a customer's behaviour indicates (partial, total or internal). Internal switching takes place when the customer changes the store but not the supermarket chain brand. Then, while internal switching also is evidence for a change, the customer still purchase in the same supermarket chain. Therefore, in global terms, the customer remains loyal to the firm.

Despite the fact that the terms 'total and partial defection' are going to be used to explain the experiment reported in this Chapter, it is important to mention that there are many other common terms found in the literature referring to defections, for example, shopping behaviour promiscuity (McGoldrick and Andree, 1997), cross-shopping (Uncles, Ehrenberg and Hammond, 1995), attritions (Van den Poel and Larivière, 2004), churns (Wei and Chiu, 2002; Larivière and Van den Poel, 2004), lost (East et al., 1995), disloyal customers (McGoldrick and Andree, 1997), customers at risk (Bloemer et al., 2002; Bloemer et al., 2003), latently dissatisfied customers (Bloemer et al., 2002; Bloemer et al., 2003), customer migration (Tao and Yeh, 2003).

5.2 Formulating the research problem

5.2.1 Introduction

The experiment carried out in this project is based on data gathered from a regional Spanish grocery chain, Supermercats Pujol, S.A (SUPSA).

SUPSA is a relatively small supermarket chain located in Lleida, Catalonia. SUPSA's leadership in this Spanish region has been achieved mainly due to its

regional expansion strategy. Each of the 70 outlets of the supermarket chain is located in the north of Catalonia. SUPSA is not a high volume retailer but is known for quality service to customers and an innovative management style.

In December 1999, SUPSA had to face a new challenge in Tremp. 5.503 people live in Tremp, which is a town in the Catalan area of Pallars Jussà. Despite the fact that the company had been the grocery leader in this town for more the 50 years, the appearance of a competitor store gave people in this town the opportunity to choose between two similar stores. Consequently, the sales of SUPSA's store suffered the new competitive situation.

In April 2000, the same competitor appeared in Flix. Flix is a town of 4.300 inhabitants where until then SUPSA was the most significant retailer. Although marketing managers would have been interested to learn from the recent experience in Tremp to avoid possible defections in Flix, there was no time to react.

From a marketing management perspective, finding the most profitable customers who were going to defect before the competitor opened the store, would have been very interesting. Armed with this information, SUPSA would have been able to take the right marketing decisions to retain its most valuable shoppers.

Identifying the customers who are going to defect is not easy. Whilst this problem is likely to be simplified by the fact that in general terms grocery shopping is considered as an habitual low involvement decision which virtually excludes cognitive problem solving process (Section 2.7.4), there are still several variables which may influence the way customers behave and make the research problem more complex. Referring to the simplification of the problem, it is important to note that future behaviours are considerably affected and explained by learnt behaviours (East, 1997). Accordingly, firms that collect, store and analyse shopping past data are more capable to forecast their regular shoppers' behaviour. As far as the complexity of the problem is concerned, based on operant conditioning learning fundamentals, purchases are explained in terms of the experienced rewards and punishments. There

are many external reinforcers (Skinner, 1953) which moderate customers' shopping behaviour. Each supermarket chain is aware of their own internal reinforcement policy, but managers do not have complete access to competitors' actions.

Moreover, SUPSA's managers were conscious that some customers defected immediately but others were likely to switch later. According to Buckinx and Van den Poel (2005), in non-contractual selling industries, it is difficult to observe when defection occurs because the majority of people do not tend suddenly to interrupt their relationship with the company. Normally, there is a period of time when customers first switch some of their purchases to another store. Evidently, marketing managers' interest would be focused not only in identifying which customers were going to switch immediately but also customers who were going to reduce their shopping baskets finally to the point of breaking their relationship with retailer. Nevertheless, there was no time to react. No extra marketing action was carried out by the marketing departments to retain customers in either of the two stores.

Despite the fact that they did not have enough time to respond with an appropriate marketing action focused to retaining the most valuable customers, they at least had the opportunity to identify each customer's behaviour, before and after the competitor opened the stores. The difference between the *'Total purchase accumulated during the 3.5 previous months'* and *'Total purchase accumulated during the 3.5 following months'* was the comparative measure selected by managers to identify customer behaviour. With this monetary value, they were able to check total defections, possible switchers and stayers. One should recall that the information will arrive too late for the retailer to react, especially to retain the valuable customers.

5.2.2 Setting Research Problem

The research problem attempts to answer these two questions:

- Which are the customers who are going to stop buying in our stores in the short term? (total defections)

- Which are the customers who are more likely to defect in the medium term?
(partial defections/uncertain customers)¹

5.2.3 Structuring the problem

The aim of the experiment is to learn from Trepmp past customers' behaviour to forecast Flix customers' behaviour. The real behaviour from the selected customers from the first store will be used to predict customer behaviour from the second store. From SUPSA's managerial perspective, Flix and Trepmp are similar villages, although they are settled down in different areas of Catalonia (See Figure 5.1).

Figure 5.1 Locations of Trepmp (Pallars Jussà) and Flix (Ribera d'Ebre)



Economically, both villages behave as the main centre of the local region. As they are in a rural area all the surrounding hinterland will come to the town not only for shopping but also for leisure and service facilities. Socio-demographic variables from each population follow the same trend as well. The population pyramid is

¹ Although the prior research problem is to finding out 'Total Defectors', it is important to note that uncertain cases are likely to be identified when interpreting the adequacy degrees of each customer to each pattern of behaviour. For instance, an uncertain customer is recognised when he/she shows a high adequacy degree in both segments (Stayers and Total Defectors) (See 5.4.3).

unbalanced. There are older people than young ones. Economically, agriculture and services are the most powerful sectors. Apart from the structural similarities, the same competitor located in both towns.

Although a deeper explanation of the methodology of the experiment is provided in section 5.4, splitting the research problem into a five stage process is needed to achieve the forecasting goal. An introduction of these 5 main stages is presented as follows:

1. Determining the behaviour of Tremp customers

It is necessary to establish a valid and agreed measure to discriminate 'Total defections' from 'Stayers' and 'Uncertains'. This criterion will be determined by the marketing managers of the company (See section 5.4.2).

2. Learning from total defections and stayers

Once the customers are classified into the 3 existing segments ('Total defections', 'Stayers' and 'Uncertain'), the uncertain will be set apart from the learning stage. Pilot tests showed that LAMDA's efficiency increases when the learning stage is based on extreme behaviours. Section 4.6.2. explains in detail LAMDA's capabilities.

3. Predicting future Flix customers behaviour

Having established the model of pattern behaviours (Tremp), the model will be used for predicting the behaviour in Flix.

4. Assignment of global adequacy degrees (GAD)

Each customer is assigned a GAD to each of the existing segments: 'Total defections' and 'Stayers'. Nevertheless, it is possible that some customers will be associated with both segments.

5. Validation of results

The purchase behaviour of customers at Flix will be classified to provide a 'control for reality'. Therefore, at this stage we will be able to compare predictions with

reality and assess the performance of the model.

5.3 Obtaining information: Identifying data sources

SUPSA, like many retailers, has been collecting data on its customers' transactions through a loyalty card programme and scanners systems. Since the introduction of the loyalty card in April 1996 across all the stores of the supermarket chain, the number of subscribers has been increasing considerably with 60.434 active cardholders now registered².

For analysis in our experiment, the 1.847 cardholders from Tresp and Flix were considered:

- 898 households registered at the SUPSA store in the village of Tresp.
- 949 households registered at the SUPSA store in the village of Flix.

It is important to mention that from the 5.503 people living in Tresp, SUPSA is presently collecting data from 1.263 loyalty card active subscribers³. In addition, from the 4.300 Flix inhabitants, SUPSA is tracking the purchasing data of 1.183 households. The loyalty card membership therefore represents a very high percentage of the households in these two small towns.

Bearing in mind that the company initiated the loyalty card programme in 1996, it is likely that they have been able to collect relevant data from the different subscribers. Part of this recorded data is going to be used to develop the experiment.

A specific data base was built for the development of this experiment. Consequently, from all the internal data gathered and stored in the company, chosen information corresponding to these 1.847 customers was considered. Moreover, this information corresponds to a particular period of time.

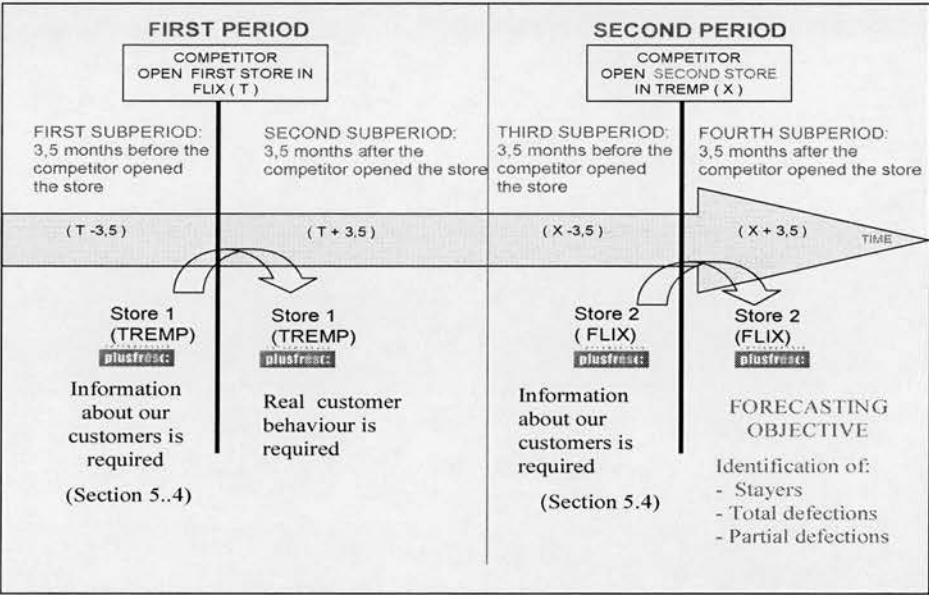
² Provided by Xavier Roure (SUPSA'S marketing manager), October 2004.

³ Provided by Xavier Roure (SUPSA'S marketing manager), October 2004.

5.3.1 Period of Observation

The Period of Observation (PoO) is the elapsed time from which the data for the experiment was taken. Therefore, as is illustrated in Figure 5.2, this experiment requires two main periods of observation.

Figure 5.2 Determination of the periods of observation



First, there is the 7 month time equivalent from September 1999 to March 2000 and the raw data was extracted from the loyalty card and scanner system database corresponding to Tremp.

Secondly, there is a second 7 month time equivalent from January 2000 to July 2000. The necessary data was taken out from the internal database corresponding to Flix.

Although both period of observation are the same length, it was known that there were differences between the sales pattern from September to March and the sales pattern between January and July. However, this dissimilarity of sales was considered and assumed away for the experiment. The opening of the competitor store affected the chosen periods of observation.

Apart from these two main periods of observation, it is important to distinguish between the first and last 3.5 months of each PoO. The main reason of this division was because the required information for the analysis would be different.

Consequently, there are four sub-periods of observation:

- (T - 3.5): From September 1999 to December 1999. This period refers to the 3,5 months before the competitor opened the store in Tremp. Information about our customers is required.
- (T + 3.5): From December 1999 to March 2000. This period refers to the 3,5 months after the competitor opened the store in Tremp. The real customer behaviour is required.
- (X- 3.5): From January 2000 to April 2000. This period refers to the 3,5 months before the competitor opened the second store in Flix. Information about our customers is required.
- (X + 3.5): From April 2000 to July 2000. This period refers to the 3,5 months after the competitor opened the second store in Flix. This information is going to be used as validating corpus.

A three and a half month time period was specifically chosen. This specific period of time was selected after an interview with the marketing managers responsible for SUPSA. Both the average of visits to each store per week and also the average shopping trips made on Saturday were known and taken into account for this decision (See Table 5.1). Experts considered that it was the minimum period of time necessary to appreciate customer potential defections.

Table 5.1 Average shopping trip frequency

	Average shopping trip (per week)	Average shopping trip on Saturdays (per week)
Tremp	1.61	0.35
Flix	1.44	0.42
TOTAL	1.52	0.38 (25%)

The first reason was to check whether Saturdays, that was the day when the company achieves its peak sales, were more frequently visited than the rest of the week. As drawn from Table 5.1, in average, 25% of the shopping trips are made on Saturdays.

The second reason was to determine how many Saturdays were required to cope with the pattern of behaviour. As is shown in Table 5.2, almost 100% of customers present a very low frequency rate of Saturdays. Note that 44% of customers do not go shopping even on one Saturday a month.

Table 5.2 Analysis of Saturday shopping trip frequency per week (SSTF)

Saturday Shopping trip frequency (SSTF) per week	Number of cardholders from Tremp within the frequency interval (out of 898)	Number of cardholders from Flix within the frequency interval (out of 949)	% of customers with this SSTF (out of 1847)
$0 < f_s < 0.25$	444	377	44,0%
$0.25 < f_s < 0.50$	218	256	25.6%
$0.50 < f_s < 0.75$	167	173	18.4%
$f_s > 0.75$	69	143	11.4%

The final decision was to select 3,5 months. We were interested to collect as many customers as possible with at least 2 shopping trips made on Saturdays, and it was the minimum time that suited our requirement. In fact, 3 months time was also a good choice. However, by checking the calendar, we discovered that some national

holidays (when the store was closed) were on a Saturday. We did not want to take the risk and lose the possibility to collect the customer behaviour on those dates. Moreover, it was known that in grocery industry partial defections are common (Buckinx and Van den Poel, 2005). Therefore, a sufficient time was required to identify at least total defections and stayers. In this case, SUPSA's Marketing manager opinion was completely taken into account: 3.5 months.

5.3.2 Selecting Data Process

Despite the fact that all the data used for the research came from company-internal customer records, a selection process was carried out mainly based on two meetings and a specific literature review.

Five people with different background and profiles joined the 2 meetings:

- A representative of SUPSA's Marketing Department,
- A representative of SUPSA's Database Department,
- A LAMDA's software tool programmer expert,
- A LAMDA's algorithm Mathematic expert,
- The researcher.

The objective of the meetings was to decide the most appropriate data to solve this problem. Therefore, it was previously necessary to know the existing data.

Despite the fact that the firm owns huge data bases related to real customer shopping behaviour, less data were available that referred to demographics and almost no data were found relating to customers' perceptions. This situation is inevitable when loyalty card data are used.

The demographic data were collected from the moment the customer registered for the loyalty card. For instance, information such as employment status, age and sex was recorded by the loyalty card application form. Moreover, information relating to the households was also collected on the loyalty card application form such as

address, pets, car, and family members. As a result, it is important to note that demographic data is normally taken once, and unless the customer expressly informs the company about any specific changes it remains unchanged.

On the other hand, behavioural data is constantly collected and reorganized. Every time a customer shops in a store of the supermarket chain, his/her purchasing behaviour is accumulated in the supermarket database. Therefore, there is a vast quantity of data. Data corresponding to the moment of shopping, the products bought, money spent, the frequency, the promotional products bought, etc.

To be more precise in the customer behaviour tracking, the company used a technology system able to collect every customer transaction, even when the customer forgets the loyalty card. Thanks to this system, when a customer does not have the loyalty card at the time of purchasing items (because he/she forgotten or because the shopper is not the cardholder of the household), the behaviour can still be tracked by the company. The customer does not lose the loyalty card advantages and the supermarket chain does not lose information from him/her. But, what is still impossible to solve is when the customer does not want to use the loyalty card, although a subscriber. In these cases, purchasing behaviour is lost.

Obviously, when the customer is not interested in obtaining the loyalty card, the company has no way to track his behaviour. Similarly there is no data when the customer buys from the competition.

From the literature review it was shown that attempting to identify customer defections in retailing is under-researched (Buckinx and Van der Poel, 2005). Nonetheless a review of the existing studies shows the different types of variables used to predict switching customers. Table 5.3 reviews publications aimed to identify customer defections in specific sectors. Table 5.4 indicates the type of predictors used to predict defections within these publications. Note that in both tables, Zeithaml, Berry and Parasuraman (1996), Buckinx and Van der Poel, (2005) and this Experiment are highlighted because of their relation between the retail sector.

Table 5.3 Forecasting defections analysis: Literature review

	Sector							Defection		Customer Base	
	Retail	Finance	Telecom	Computer Manufacturer	Insurance	Automotive firm	Other service	Total	Partial	Complete	Partial
Bloemer et al (2003)		x						x	x	x	
Athanassopoulos (2000)		x						x		x	
Bhattacharya (1998)								x		x	
Keaveney and Parthasarathy (2001)							x	x		x	
Lemon et al. (2002)											
Mittal and Kamakura (2001)						x	x	x		x	
Mozer et al. (2000)			x					x		x	
Popkowski et al (2000)								x		x	
Van den Poel and Lariviere (2004)		x						x		x	
Weerahandi and Moitra (1995)			x					x		x	
Zeithaml et al. (1996)	x			x	x			x		x	
Buckinx and Van der Poel (2004)	x								x		x
This Experiment	x							x	x	x	

Source: Adapted from Buckinx, W. and Van den Poel, D. “Customer base analysis: partial defection of behaviorally loyal clients in a non-contractual FMCG retail setting”. *European Journal of Operational Research*, 2005, 164 (July), 252-268.

SUPSA’s loyalty card programme is not able to gather customer perceptions (opinions, emotions) but enables collection of behavioural data. Despite the fact that the company developed means of knowing customers’ opinions, such as specific satisfaction surveys or a customer call centre for customer complaints, there is no easy and profitable way to relate these subjective information sources coming from some customers with the objective behavioural shopping information coming from loyalty card use. For these reasons, the predictors for our experiment will be chosen from the data accumulated from each cardholder, which is behavioural and demographic data (in particular household traits) (See Table 5.4).

Table 5.4 Predictors used to predict customer defections

	Behavioural Antecedents	Socio-demographics	Perceptions
Bloemer et al (2003)		x	x
Athanassopoulos (2000)		x	x
Bhattacharya (1998)	x	x	
Keaveney and Parthasarathy (2001)	x	x	x
Lemon, White and Winer (2002)	x	x	x
Mittal and Kamakura (2001)		x	x
Mozer et al. (2000)	x	x	
Popkowski, Sinha and Timmermans (2000)	x	x	
Weerahandi and Moitra (1995)		x	
Zeithaml, Berry and Parasuraman (1996)			x
Buckinx and Van den Poel (2005)	x	x	
This Experiment	x	x⁴	

Source: Adapted from Buckinx, W. and Van den Poel, D. "Customer base analysis: partial defection of behaviorally loyal clients in a non-contractual FMCG retail setting". *European Journal of Operational Research*, 2005, 164 (July), 252-268.

5.3.3 Predictors of defection used in this study

In total, 38 variables were initially selected for building the specific database for the experiment. These variables can be split into two main groups: socio-demographic and behavioural.

Socio demographic details

Customer demographics have been extensively applied to predict customer defections (See Table 5.4). Consequently, several demographic predictors available in the internal database are selected. The *socio demographic* group of variables from the experiment includes 10 variables. As it is shown in Table 5.5, the variables are related to the *cardholder subscriber* (employment status, gender, age) and also the details related to his/her *household* (family members, loyalty, potentiality, animals, length of relationship with supermarket, ownership of a washing machine).

⁴ The socio demographic traits are directly related to the household, not to the personal individual traits.

Table 5.5 Socio demographic predictor variables selected for the experiment

N	Variable name	Description
Information related to the individual Cardholder		
1	Customer code	The loyalty card code which allows to identify the customer
2	Employment Status	There is a classification between: house woman, retired, unemployed, employed, employee, self-employed
3	Gender	Women/Men
4	Age	Date of Birth
Information related to Household		
5	Family members	Household size: number of members in the household
6	Loyalty	Specific measure by SUPSA. Larger families need to buy larger quantities of products due to their higher level of consumption.
7	Potentiality	The potentiality measures how much capacity of purchase this household has, according to the normal basket of SUPSA's single customer.
8	Animals	Presence of pets: yes/no
9	Length of relationship with supermarket	Time between the loyalty card subscription and the last current purchase with the company
10	Ownership of a washing machine	Indirect variable related to income : Not having a washing machine was considered 'low' income.

Despite the fact that the cardholder subscriber plays an important role (buyer) in the purchasing decision process, and his/her age, gender and employment status may reflect the stage of the family cycle, it is important to note that his/her purchased basket is considerably influenced by the rest of the members in the household. Therefore, information related to the household it is also essential:

Family members, larger families need to buy larger quantities due to their higher level of consumption. They also require a greater variety of products (Bawa and Ghosh, 1999).

Household income is likely to be positively related to expenditure per trip (Bawa and Ghosh, 1999) and also to the purchased brands. As SUPSA did not have direct information for income, a variable was initially chosen as indirect information

providers of 'household income'. This variable was *the ownership of a washing machine*. Determining a variable which gave indirect information about the income was one of the first objectives of the meetings and it was not easy. After presenting several alternatives, the assumption was agreed that not having a washing machine indicated low income. Obviously, 'Do you have a washing machine?' was not a question asked in the loyalty card subscription. However, we were able to know the answer just reviewing which households had purchased, at some time some washing machine specific product.

Loyalty and Potentiality are two specific measures provided by the supermarket chain. Both variables include 4 categories: A, B, C and D (See Table 5.6).

Potentiality is a variable considered as a qualitative function depending on the number of family members. A household of 1 individual has a 'D' in potentiality, while a household of more than 4 members is given an 'A' in potentiality.

$$P(customer) = F (number\ of\ family\ members)$$

Loyalty is measured by SUPSA as a qualitative function according to the annual purchases made by a household divided by the annual amount of money spent by a Spanish individual. This second number, which is provided by El Instituto Nacional de Estadística (INE), is 200.000 Ptas., (1.200 euros).

$$L = F (Annual\ total\ purchases / 200.000Ptas)$$

It is important to note that these variables are interpreted together (See Table 5.6). The first letter corresponds to 'loyalty' and the second letter to 'potentiality'. For example, a cardholder classified in 'DA' group shows that despite the fact that they are a big family; their purchases in the supermarket are very low. In the opposite way, an AD customer relates to a customer who spends the maximum in our supermarket but is single.

Table 5.6 Measuring loyalty and potentiality

		A	B	C	D
		>4	3-4	2	1
Potentiality					
Family members					
Loyalty		>875.000	401.000-875.000	201.000-400.000	<200.000
A	>75%	AA	AB	AC	AD
B	50%-75%	BA	BB	BC	BD
C	25%-50%	CA	CB	CC	CD
D	< 25%	DA	DB	DC	DD

In supermarkets, customers seem to switch relatively often (Roos, 1999), so *length of relationship* is used to determine not only whether the customer is a ‘veteran’ to the company but also the evolution of the purchasing pattern. Although this variable was selected by the members of the meeting, it was rejected. We realised that most of the customers had the same length of relationship. The reason was that when the company launched the loyalty card programme, the pilot test was developed in these two villages. Under these circumstances, the information provided by this variable was not very relevant for the experiment.

Behavioural antecedents

As it is listed in Table 5.7, the *Behavioural Antecedents* includes 23 predictors classified into various types of variable: Shopping trip frequency, Promotional behaviours, Shopping behaviour across product categories, Brand purchase behaviour and Monetary indicators.

Table 5.7 Behaviour predictor variables selected for the experiment

N	Variable name	Description
Shopping Trip Frequency		
11	Mean No. of days per week on which purchases are made	Average number of trips to the shop by week
12	N° of shopping trip	Number of shopping trips (total PoO)
13	% Purchases made on Saturdays (number of customer trips)	Percentage of monetary value made on Saturday (from the total shopping trips made during the PoO)
14	% days tinned purchased / Total number of shopping days in period	Percentage of days that customer buys tinned products from the total trips in the PoO
15	% days meat purchased / Total number of shopping days in period	Percentage of days that customer buys meat from the total trips in the PoO
16	% days pre-packed meat purchased / Total shopping days in period	Average of days that customer buys pre-packed meat from the total trips in the PoO
17	% days vegetables purchased / Total shopping days in period	Percentage of days that customer buys vegetables from the total trips in the PoO
18	% days frozen goods purchased / Total shopping days in period	Percentage of days that customer buys frozen goods from the total trips in the PoO
19	% days bread purchased / Total shopping days in period	Percentage of days that customer buys bread from the total trips in the PoO
Promotional Behaviours		
20	% of purchases made up by special offers	Percentage monetary value of promotions (per total purchases made during the PoO)
21	1% sales destination	SUPSA provides the customer the opportunity to spend the 1% of his sales in products, charity or money
Shopping Behaviour across Product Categories		
22	No. of different articles purchased in the period.	Number of different items purchased during the PoO
23	No. of departments where no purchases were made.	Number of departments where NO item was bought during the PoO
24	% purchase of tinned goods /Total purchases	Percentage of monetary value spend on tinned goods from the total items purchased
25	% meat purchases / Total purchases	Percentage of monetary value spend on meat from the total items purchased
26	% purchase of pre-packed meat / Total purchases	Percentage of monetary value spend on pre-packed meat from the total items purchased
27	% purchase of vegetables /Total purchases	Percentage of monetary value spend on vegetables from the total items purchased
28	% purchase of frozen goods / Total purchases	Percentage of monetary value spend on frozen goods from the total items purchased
29	% bread purchased /Total purchases	Percentage of monetary value spend on bread from the total items purchased
Brand Purchase Behaviour		
30	% purchase of supermarket's private label brand	Percentage of monetary value spend on supermarket's private label brand from the total purchases
31	Purchase of Brand Products (amounts)	Total monetary amount of spending on Brand Products
Monetary Indicators		
32	% Purchases made on Saturdays (amount)	Percentage of monetary value spend on Saturdays (from the Total Purchase of PoO)
33	Total Purchases	Total monetary amount of spending during PoO

Shopping Trip Frequency, this type of variable includes different versions of the frequency variable. Frequency is the number of shop visits (Buckinx and Van den Poel, 2005). Apart from the average number of trips to the shop by week, variables explaining the frequency when categories are purchased during the PoO are also included. According to Buckinx and Van den Poel (2005:9),

'Customer's frequency of purchases may be predictive for their future behaviour'.

Promotions, for some customers, lower prices are the only explanation for their purchases. Consequently, people being more sensitive to promotions are likely to have a higher trend to store switching and thus defection (Buckinx and Van den Poel, 2005). *Percentage of purchases made up by special offers* and *1% destinations* will provide customer price sensitivity. 1% destinations variable corresponds to the reward alternatives that SUPSA gives to its customers. Customers can choose between obtaining a reduction corresponding to the 1% of the total purchases made by him within a year, or donating this money to the Church, or donating the corresponding amount to a non-governmental organization.

Category, defection may occur when customers are not pleased with a specific product or service (Rust and Zahorik, 1993). Consequently, a list of variables representing the spending in each category of the retailer was selected. In addition, the variable which shows the number of departments (categories established by SUPSA) where no purchase was made during the PoO was also selected as a defection predictor.

Brand, SUPSA's product policy classifies each item into two main brand categories: Supplier brand (there is no distinction between national or international) and Private label brand (they do not have a retailer own brand). For each of these categories a variable is compiled, representing the relative spending of a customer. According to Corstjens and Lal (2000), retailer brands (private label or own store brand) can be a powerful tool to differentiate a store from its competitor and increase store loyalty. Therefore, following Corstjens and Lal's statement, it seems that the more a

customer spends on private label brands, the lower likelihood of this customer switching to a competitor.

Monetary, past behavioural variables, more specifically RFM (recency, frequency and monetary value) variables are the best predictors of partial customer defection (Buckinx and Van den Poel, 2005). Monetary value indicators represent the amount of money someone has spent at the company. Despite the fact that most of the variables included in the preceding group *product categories* also provide information about customer spending, it was decided to add this new group called *Monetary* for representing the amount of money a customer has spent at the company during the PoO.

Predicted Variable: 'Defection' behaviour

Interpurchase variations, this type of behavioural variable not only comprises the *total purchases in the 1st period* and the *total purchases in the 2nd period*, but also the monetary (Ptas.) and percentage variations between them (See Table 5.8).

Two important points are relevant at this stage.

- Firstly, in both stores, customers' behaviour before and after the entrance of the new competitor was known, so the company already knew which customers defected and which remained with the chain (*situation after competitor opened the store, V38*) . Therefore, the real behaviour of the customers from Tresp was set apart until the control stage.
- Secondly, SUPSA's managers stated that there was no customised marketing campaign focused on retaining the most valuable customers carried out by the company in any case, in Flix or Tresp. Based on that, any bias from a possible marketing action would influence the final predictions was avoided.

Table 5.8 Interpurchase variations predictors and the predictive variable

N	Variable name	Description
Interpurchase Variations		
34	Total purchases in the 1st period	The accumulated purchases (ptas) referred to the 3.5 months previous to the appearance of the competitor store
35	Total purchases in the 2nd period	The accumulated purchases (ptas) referred to the 3.5 months after the appearance of the competitor store
36	Variation (in Pesetas)	The difference between both spending periods
37	Variation (in %)	The percentage of the difference between both spending periods
Predictive Variable		
38	Situation after opening of competitor's outlet	The label chosen by the members of the meeting to classify behaviours: Stayers, Total defectors, uncertain.

Variable 38 is the predictive variable. It is based on the results extracted from the variable 37 (*Variation in %*). Variable 37 is the percentage resulting from comparing the accumulated purchases spent by the same customer during the 3.5 months previously to the moment that the competitor opened the store in Tresp (V34) and the purchases made during the 3.5 months afterwards (V35), based on the accumulated purchases spent during the first 3.5 months.

Once the percentages were known, a ranking of all the 884⁵ customers from Tresp, ordered by V37 was made. Observing the ranking, three different patterns of behaviours were found by the membership of the meeting. It was a commonly shared judgment that all the customers whose purchases had increased dramatically ⁶would be labelled as 'Stayers' and classified in group 2. All the customers whose purchases had decreased dramatically would be classified in group 1 (Total defectors). The rest of the customers were considered 'Uncertain' cases, and they were assigned to group 3. The landmarks (percentages) which are going to be used to delimit and label the different segments are defined in section 5.4.2.

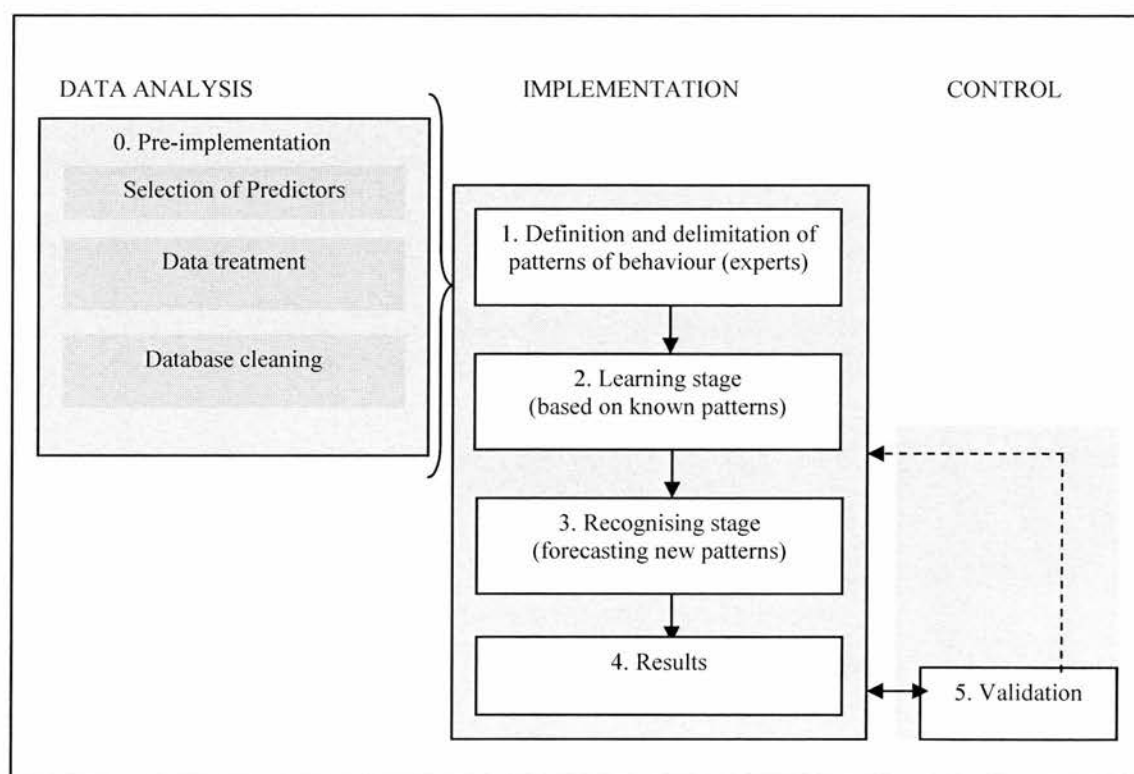
⁵ The initial list of 898 individuals was reduced to 884. 15 customers had lots of missing values and were excluded from the analysis.

⁶ Table 5.12 illustrates the percentages (landmarks) which delimits and defines a dramatic increase or dramatic decrease.

5.4 Implementation of the forecasting model

The forecasting model based on known behavioural patterns introduced in this section follows 3 main phases. As illustrated in Figure 5.3, the first phase is the data analysis. Data analysis is required prior to the implementation. At this stage, the final predictors are chosen, the data treatment and cleaning carried out. Once this pre-implementation stage is finished, the implementation itself is started. The implementation phase is composed of 4 sequential stages which includes pattern behaviour delimitation, learning, recognition and results. Once the results are known, the final validation phase is then developed as the control phase.

Figure 5.3 Forecasting model based on known behavioural patterns



5.4.1 Data analysis

The data analysis procedure aims at achieving the following proposals:

- To select the most valuable predictors for the research goal of the experiment.

- To know whether a transformation process of the variables is required.
- To clean the customer database.

The first objective of analysing the data was to finding out whether the initial selected predictors were significant enough for the research problem. Despite the fact that the significance of each variable would have been statistically measured (For instance, Chi-Square analysis for qualitative variables and ANOVA for the quantitative variables), expert’s judgements were finally considered. As explained in Chapter 4, not all the types of information have a rational and clear explanation. Moreover, if existing, this explanation is not easy to justify or quantify because it is mainly based on common sense and past experience. Related to that, when deciding on the final predictors, only expert’s judgements were considered. From the initial 38 variables, just 31 of them were selected and analysed. Table 5.9 illustrates the variables that were eliminated and summarises its reasons.

Table 5.9 Ruled out predictors

Eliminated predictors	Judgemental reason
<ul style="list-style-type: none"> - Information related to the individual Cardholder <ul style="list-style-type: none"> Employment status Gender Age 	<ul style="list-style-type: none"> - Household was the subject of research - The individual traits from the cardholder (buyer) did not explain household shopping patterns
<ul style="list-style-type: none"> - Information related to Household <ul style="list-style-type: none"> Animals Length of relationship with supermarket Ownership of a washing machine 	<ul style="list-style-type: none"> - Animal not describes defect - Length may help to anticipate defections but all the cardholders showed the same. - Ownership washing machine not describes possible defections

The cardholder specific information was set apart because the subject of research was the household⁷ and not the buyer (as the person in charge of the purchase). Buyers’ socio demographic information such as employment status, gender or age was not taken into account when explaining household behavioural patterns.

Furthermore, although the household is the central of research, there were 3 related variables that were not considered by the experts when attempting to anticipate

⁷ From now on, the term customer behaviour will also refer to household behaviour.

customer defections. These variables are listed in Table 5.9. As far as ‘animals’ predictor is concerned, there were so many incomplete fields that makes the predictor useless. Then, it was ruled out. The ‘ownership of a washing machine’ was affirmative in every customer register. The information was valuable. Related to the ‘length of the relationship with the supermarket’ was also the same in every case, as the pilot test when the company launched the loyalty card programme was carried out in Tremp.

The second objective was to decide whether a transformation process was necessary for some of the variables. For instance, qualitative variables which show a low number of observations in one of their categories, the category with lack of observations were merged to another one. In the same way, a transformation of some quantitative variables was carried out as well. This specific transformation is called discretization. Discretization takes place when a numerical variable is partitioned into a number of sub-ranges and each sub-range is treated as an ordered category (order of magnitude). Although it may result in a possible loss of information, a simplification of information (but not loss) is achieved with discretization (Agell et al., 2005). Figure 5.4 is an example of a quantitative variable and Figure 5.5 illustrates the same variable being transformed as an order of magnitude.

Figure 5.4 Example of quantitative variable:

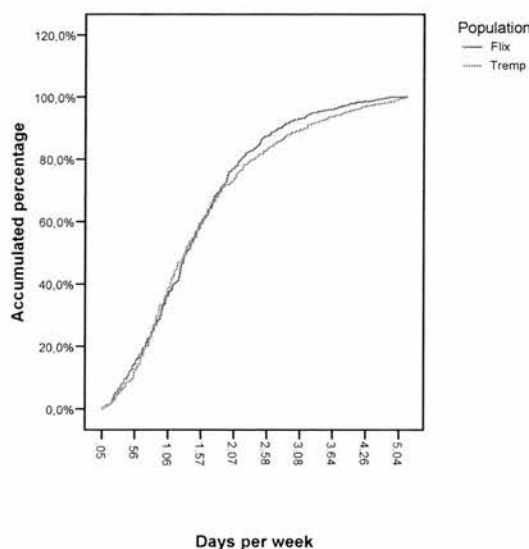
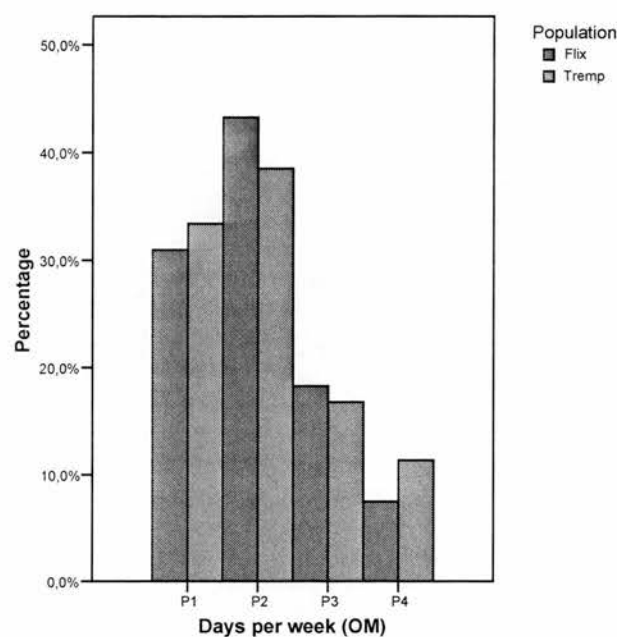


Figure 5.5 Example of the order of magnitude variable



The process of discretization was carried out with all the predictors included in the mentioned group of Shopping Frequency Trip (from Variable 11 to Variable 19). Pilot tests were previously required to decide whether it was better to transform each numerical variable or not. The reason why the discretization process was applied to the shopping frequency variables was that the basic metric of all these variables was days and trips. Then, the decision that this basic metric should be better interpreted as qualitative than quantitative was taken. Also number of family members was transformed. The reason was that the basic unit (persons) was better interpreted when it took an ordinal value rather than a quantitative value.

Consequently, the final predictors were classified. The predictors encompass both qualitative and quantitative data. As illustrated in Table 5.10, 11 of the total variables are qualitative and 19 quantitative. Nine qualitative variables have been partitioned. Initially, all the predictors which present orders of magnitude (OM) were quantitative variables.

Table 5.10 Predictors classification according to its type of data: qualitative (non-numerical) and quantitative (numerical)

Qualitative		Quantitative
Being discretized	Predictor	Predictor
(OM)	Mean No. of days per week on which purchases are made	% of purchases made up by special offers
(OM)	N° of shopping trip	1% sales destination
(OM)	% Purchases made on Saturdays (number of customer trips)	No. of different articles purchased in the period.
(OM)	% days tinned purchased / Total number of shopping days in period	No. of departments where no purchases were made.
(OM)	% days meat purchased / Total number of shopping days in period	% purchase of tinned goods / Total purchases
(OM)	% days pre-packed meat purchased / Total shopping days in period	% meat purchases / Total purchases
(OM)	% days vegetables purchased / Total shopping days in period	% purchase of pre-packed meat / Total purchases
(OM)	% days frozen goods purchased / Total shopping days in period	% purchase of vegetables / Total purchases
(OM)	% days bread purchased / Total shopping days in period	% purchase of frozen goods / Total purchases
	Loyalty/potentiality	% bread purchased / Total purchases
(OM)	Family members	% purchase of supermarket's private label brand
		Purchase of Brand Products (amounts)
		% Purchases made on Saturdays (amount)
		Total Purchases
		Customer code
		Total purchases in the 1st period
		Total purchases in the 2nd period
		Variation (in Pesetas)
		Variation (in %)
TOTAL	11 qualitative	19 quantitative

The third step was to clean up both Tresp and Flix customers' database. It is important to note that 15 customers were removed from Tresp and 161 from Flix. Customer records that had many missing values were not valuable so they were removed.

5.4.2 Implementing LAMDA supervised forecasting model

Definition of behaviour patterns landmarks

As was previously explained in Chapter 4, LAMDA can be used to learn

‘intelligently’ from previous shopping behaviour patterns to identify future individual behaviour, simultaneously integrating qualitative and quantitative data. However, the previous shopping behaviour patterns need to be defined and delimited before taking any next step.

Despite the fact that the behaviour of each customer from Tresp was known (See Table 5.11), and the variation of purchases was also collected, defining the boundaries which showed the situation of each customer after the competitor opening was not so easy.

Table 5.11 Example of variations in each customer behaviour (before and after competitor’s store opened)

Identification Code	Total Purchase in the first period	Total Purchase in the second period	Variations (in pts.)	Variations (in %)	Situation after competitor opening
9999000005218	290,863	250,118	-40,745	-14%	?
9999000010991	156,839	33,663	-123,176	-79%	?
9999000010052	25,296	38,654	13,358	60%	?
9999000010106	70,362	70,746	384	1%	?
9999000010298	129,413	145,780	16,367	13%	?
9999000010434	22,296	35,654	13,358	60%	?
9999000010595	61,108	96,444	35,336	58%	?
9999000010861	94,856	14,361	-80,495	-85%	?
9999000010885	129,714	69,529	-60,185	-46%	?
9999000010922	87,748	43,545	-44,203	-50%	?
9999000010960	242,441	85,515	-156,926	-65%	?

Taking the variations (in %) into account, some confusion arose between the experts. For instance, some experts in the Marketing department disagreed with the fact that a decrease in 14% of the purchase meant a customer loss. Also there was disagreement on whether an increase of 1% showed a clear ‘Stayer’. However, a common agreement was reached in accepting that a 79% decrease in purchases showed a total defector. Also a 60% increase confirmed a clear ‘Stayer’. We realised that there was a subjective interpretation of the percentages which created some confusion when determining the boundaries of the ‘Total defectors’ and ‘Stayers’, but this confusion just appeared in the uncertain cases, not in the extreme. Based on that, a percentile

division was developed.

Table 5.12 Quartile distribution from Variable 37 (Variations %)

	Number of customers from Tresp joined	Landmark
4 th quartile: STAYERS	221	+17.97%
3 rd quartile	221	+4.67%
2 nd quartile	221	-51.03%
1 st quartile: TOTAL DEFECTORS	221	-98.66%

Then, the problem of deciding the numerical boundaries which distinguished between patterns of behaviour (Stayers, Uncertain and Total Defectors) was solved. As illustrated in Table 5.12, the 25% of customers from Tresp with the highest positive variations were joined in the same quartile. 221 customers from the 4th quartile were labelled as 'Stayers' and assigned to segment 2. On the other hand, the first quartile joined the 25% customers from Tresp with the highest negative variations, from 51,03% to 98.66% decrease. These 221 customers were labelled as 'Total defectors' and assigned to segment 1. The second and third quartile joined the rest of the customers (Uncertain: Group 3). Table 5.13 illustrates an example of the customer's classification according to its location in the quartiles division.

Table 5.13 Example of the customers' classification according to the purchase variation and quartiles division

Identification Code	Total Purchase in the first period	Total Purchase in the second period	Variations (in pts.)	Variations (in %)	Situation after competitor opening
9999000005218	290,863	250,118	-40,745	-14%	Uncertain (3)
9999000010991	156,839	33,663	-123,176	-79%	Total defector (1)
9999000010052	25,296	38,654	13,358	60%	Stayer (2)
9999000010106	70,362	70,746	384	1%	Uncertain (3)
9999000010298	129,413	145,780	16,367	13%	Stayer (2)
9999000010434	22,296	35,654	13,358	60%	Stayer (2)
9999000010595	61,108	96,444	35,336	58%	Stayer (2)
9999000010861	94,856	14,361	-80,495	-85%	Total defector (1)
9999000010885	129,714	69,529	-60,185	-46%	Uncertain (3)
9999000010922	87,748	43,545	-44,203	-50%	Uncertain (3)
9999000010960	242,441	85,515	-156,926	-65%	Total defector (1)

Learning stage

For the learning process, just the extreme customers were used. As in humans, learning becomes easier when the learning task is repetitive and when there is a wide difference between opposite cases. Accordingly, the greater the difference present in the known patterns of behaviour, the easier it is to learn from them. In the light of this, 442 customers were chosen for the learning stage. These 442 customers, coming from the 1st and 4th quartile were called learning corpus (LC). These customers from these two quartiles were chosen because they presented the most different behaviour.

Also in this stage, the degree of tolerance and the fuzzy connective which better fits with the introduced data was selected, according to distributions of the chosen variables. The selected fuzzy connective was Minmax algorithm.

Recognising stage

Once behaviour patterns from the LC were learned, the recognition stage began. For the recognition stage, information coming from customers from Flix (recognising corpus) was required. Each customer contained the same predictors as the learning corpus except the group of variables related to inter-purchase variation group (V35, V36 and V37) and the predicted variable (situation after competitor aperture, V38). Based on the connective learning and the tolerance degree determined in the learning stage, each customer from the recognising corpus (RC) was assigned to each segment, ('Stayers' and 'Total defectors') with an adequacy degree. Then, the forecasting task in strict sense was carried out in this stage. The model anticipated the behaviour from the RC accordingly to the behaviour learned from the LC. Then, the customers who were likely to defect and the customers who were likely to stay when the competitor opened the new store are not only recognised but also identified.

Results

The results obtained were analysed from a marketing standpoint (See Figure 5.6). The main goal was to predict which customers from Flix would keep shopping at the

store in spite of the entry of a new competitor, and which ones would simply stop buying at the SUPSA’s outlet. If the predictions were accurate, the company would focus its efforts on the total defectors who really needed special attention.

Figure 5.6 Interpretation of results based on a marketing standpoint

		LAMDA’s model forecasts	
Real behaviour		Stayer	Total Defector
	Stayer	SUCCESS	ERROR
	Total Defector	ERROR	SUCCESS

The consequences of wrongly allocating a member in the ‘Total Defectors’ group as belonging to the ‘Stayer’ are that SUPSA would not invest in retaining that customer. On the other hand, customers from the ‘Stayer’ segment do not need additional attention. However, if they were erroneously classified in the ‘Total Defector’, SUPSA would unnecessarily spend money on satisfied customers. Thus, the most important goal was to identify the customers who were likely to defect. Therefore, if one had to commit a mistake, it was preferable to assign more individuals to this class (thus producing false alarms) rather than fewer.

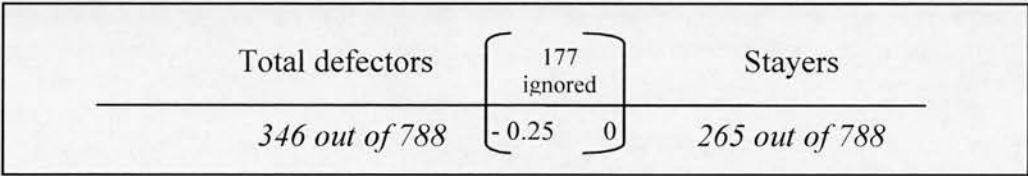
5.4.3 Validation

This is the control stage. In this stage, variables which had moved apart in the recognising stage (V35, V36, V37 and V38) were essential at this point as LAMDA’s forecasts were checked with real behaviours. As mentioned before, although the data related to the real behaviour was not used for the experiment, it was crucial for testing the validity of the model.

But, prior to the validation itself, professional experts’ opinions were again required. Establishing the values that indicate the ‘Stayers’ and the ‘Total Defectors’ was needed. Although the rankings of the interpurchase variations (%) were known, experts had to define Flix’s boundaries for stayers and for defectors.

In this stage, finding the extremes was not as important as reaching a common view and internal agreement of what ‘Stayer’ and ‘Total defector’ were. Based on that, two assumptions were established. Firstly, the firm would assume that all the variations higher than 0% were considered ‘Stayers’. Secondly, the variations lower than minus 25% were considered ‘Total Defectors’. The rest would be treated as uncertain but ignored for the results matching.

Figure 5.7 Definition of the boundaries for the validation corpus (VC)



Based on that, the validation corpus (VC) comprised 346 customers who presented more than 25% variation decrease in their purchases (Total Defectors) and by the ‘Stayers’, 265 customers who showed a 0% or more variation in their purchases. The rest, 177 were considered uncertain cases and ignored for the validation matching (See Figure 5.7).

Once the validation corpus was decided, the matching between the model’s forecasts and real behaviour was developed. It is important to remember that LAMDA deals with both multiple global adequacy degrees (MGAD) and maximum adequacy degree (MaxGAD). Then two possible interpretations may arise.

Table 5.14 Interpretation of results based on adequacy results

	Identification code (Flix)	GAD in segment ‘Stayers’	GAD in segment ‘Total defectors’	Maximum GAD (MaxGAD)	Prediction Based on MGADs	Prediction based on Maximum
a)	9999000019031	0.47	0.09	0.47	Stayer	Stayer
b)	9999000020153	0.04	0.03	0.04	Uncertain	Stayer
c)	9999000001537	0.78	0.80	0.8	Uncertain	Total Defector
d)	9999000080273	0.2	0.67	0.67	Total Defector	Total defector

As illustrated in Table 5.14, when interpreting the results under a MaxGAD standpoint, a customer is assigned to the segment which has obtained the highest GAD. However, when interpreting the results according to multiple adequacy degrees (MGAD), the ‘uncertain’ segment appears. Uncertain is the result of not perceiving sufficient difference between the GADs. For example, customer from row b) presents a low GAD in both segments. The difference between 0.04 and 0.03 is virtually insignificant. Based on that, although the maximum GAD interpretation would assign this customer to ‘Stayer’ segment, in reality, the behaviour is not so clear, and the customer is located to the ‘Uncertain’ segment as future behaviour is not clear. Customer represented by row c) is considered uncertain as well. He/she presents a high GAD in both segments so he/she has the possibility to behave as ‘Stayer’ and as ‘Total Defector’. Therefore, she/he is considered and located to the ‘Uncertain’ segment again. Based on that, validation stage is analysed considering both options, the maximum adequacy degree and the multiple GADs:

Maximum adequacy degree is considered (no overlapping)

The customer is located to the segment which has the highest GAD. Based on that, the following matrix is resulted:

Table 5.15 Interpretation of numerical results based on Maximum adequacy degree

		LANDA’s model forecasts	
Real behaviour		Stayer	Total Defector
	Stayer	205	60
	Total Defector	240	106

The forecasting success percentage is 50.9%. From 688 customers the model was able to forecast successfully 205 (Stayers) and 106 (Total defectors).

Multiple GADs are considered (overlapping)

This capability of the model to assign each customer to each pattern of behaviour (segment) has a managerial advantage which consists of interpreting the results by a not crisp point of view. That means that an overlapping between segments is accepted.

In particular, for validating the results based on these overlapping of behaviours, customers with similar adequacy degree in both segments are assumed to belong to both segments simultaneously. It is important to note that a similar adequacy degree is defined as the difference between the GADs in each segment and it has been assumed to be 0,015 or less. Taking this simultaneity into consideration, a new matrix resulted (See Table 5.16).

Table 5.16 Interpretation of numerical results based on multiple GADS

		LAMDA's model forecasts	
Real behaviour		Stayer	Total Defector
	Stayer	249	178
	Total Defector	322	236

From the 346 customers who really defected, the model anticipated 236 customers. Then the success ratio is 68.21%. The model was not able to anticipate 110 out of the total defectors. This is an error of the 31.79%. From a marketing standpoint, this 31.79% is a major error as the firm would take no steps to retain these customers, as the firm would assume that these customers would remain purchasing in their supermarket. From the 265 stayers, the model was able to correctly forecast 249 behaviours. Then, a 93.96% of the stayers have been recognised. However, there were 6.04% of the real stayers who were identified as defectors. From a marketing standpoint, these customers should receive a special marketing action to retain them. The cost of the marketing action would have been saved as the customers were stayers. Only a minor error is evident in this case.

Summary

A forecasting model based on known behavioural patterns is the basis of the experiment. Although implementation is the essential part of the model, the previous analysis of data is definitely recommended as it may simplify the research.

The experiment attempts to forecast the customers who are likely to defect when the competitor opens a new store. Based on historic shopping data, the fuzzy connective algorithm learns from past patterns of behaviours to anticipate future behaviours in similar situations. Although the learning and recognition stages are automatically performed, a professional expert opinion is essentially required when delimiting the boundaries between the existing patterns of behaviour. Also when defining the validating corpus boundaries.

As in humans learning, the learning stage is better performed when the inputs are clear, real and different.

- Clear means as unambiguous as possible. The more extreme the inputs behaviours, the more effectively LAMDA learns from them.
- Real means known. There is no space for judgements. Once the patterns are delimited, they became the existing patterns which represent the reality.
- The more different the input segments are (patterns of behaviour), the most likely to be successful with the predictions the model is as it leans from extreme cases, so it is easier.

In the recognition stage, the identification and forecasts of behaviour is developed. When identifying the customers, each customer is located to the segment which presents the maximum adequacy degree. However, a multiple assignation of the same customer to more than one segment is also possible by dealing with the multiple GADs provided by LAMDA's approach. Then, each customer is located to each existing pattern of behaviour with an adequacy degree.

When validating the results, higher success rate is achieved when overlapping is

accepted rather than when assigning a customer in just one segment. Customers' behaviour is not always black or white. When considering the possibility that a customer may behave in different ways, the forecasts are more successful. At this point it is important to note that the major interest was to find out the customers who were going to defect. Therefore, although LAMDA was able to recognise the 93.96% of the 'Stayers', this ratio was not as interesting as the forecasting success ratio of 'Total Defectors'. Mentioned that, noting the forecasting success rate of 68.21% which is achieved when the multiple GADs are considered.

CHAPTER 6

Experiment 2: Forecasting customer behaviour in the Spanish grocery industry: Identifying the customers who are going to buy online

6.1 General research lines

Predicting and understanding online-buying behaviour is of major importance for e-commerce website managers (Buclin and Sismeiro, 2003). Within the research literature, there are several attempts to support and facilitate the achievement of this managerial goal.

Basically, researchers in this area attempt to answer these two main questions:

- Which are the predictors of consumer's online purchasing?
- Is it possible to predict and explain customers channel choice between store and Website?

Which are the predictors of consumer's online purchasing?

Different types of indicators have been proposed to predict online purchase. They can be classified into 5 main groups:

- Socio demographic characteristics,
- Internet shopping positive attributes,
- Perceived transaction costs,
- Online historical behaviour and
- Site design.

The first group emphasise the importance of considering *socio-demographic* characteristics (Shim and Drake, 1990; Breitenbach and Van Doren, 1998; Douthu and Garcia, 1999; Mathwick, Malhotra and Rigdon, 2002; Raijas, 2002; Van den Poel and Buckinx, 2005). Despite the fact that there is not a perfect match between the profiles presented in the different reports, age, gender, education degree and

income are distinct variables to describe Internet users' profile. In addition to these socio-demographic features, race and language are also proposed by Padmanabhan, Zheng and Kimbrough (2001). Then, it seems that demographic information contributes to classifying customers as online and non-online buyers (Van den Poel and Buckinx, 2005).

The retailing research literature also suggests that consumer perceptions are important indicators of the probability of making purchasing decisions on the Internet. Usually, this type of primary data (perceptions) is collected from quantitative surveys. A list of positive *attributes of internet shopping* such as time saving, money saving, easy accessibility, simple interface screen, accessibility to select of alternatives and interactivity is provided by Kim and Kim (2004). On time delivery, customer service, privacy policies and shipping and handling are also positive attributes provided by Reibsten (2002). There is an evident interest in finding the appropriate attributes required to predict the propensity to buy online within the literature (e.g. Zeithaml and Gilly, 1987; Breitenbach and Van Doren, 1998; Vellido, Lisboa and Meehan, 1999; Crawford, 2000; Shim, Eastlick and Lotz, 2000; Ray, 2001; Chiger, 2001; Supphellen and Nysveeb, 2001; Hansen, Jensen and Solgaard, 2004).

In addition to the previous two groups, consumer's perceptions of *internet transaction costs* have been reported to predict consumers' intentions to buy products or services (e.g. Cohen, 1990; Shim, Eastlick and Lotz, 2000; Rajas, 2002; Gupta, Su and Walter, 2004; Chiang, Zhang and Zhou, 2004). For example, based on transaction cost analysis framework, Chiang, Zhang and Zhou (2004), distinguish between 7 types of possible costs when assessing a customer's attitude toward web shopping. These seven are search, comparison, examination, opportunity, payment, delivery and post service costs. According to Bakos (1997), reducing these costs would be interesting for web managers as it would help to enhance the website information and monetary flows.

There is research which provides forecasts of online purchasing based on captured

online historical behaviour, mainly known as ‘clickstream’ data. Conversely to the physical stores, this approach is focused on every step that the visitor takes in the Website.

‘Clickstream data, typically contain the trajectory of (prospective) clients at the company’s website’ (Van den Poel and Buckinx, 2005: 1).

Then, based on this tracking, researchers are able to know the specific visitor who enters to the website but does not buy. Interest in analysing the conversion of store visits into purchases based on historical online visiting data is found over the literature (Moe and Fader, 2001; Moe and Fader, 2002; Buclin and Sismeiro, 2003; Van den Poel and Buckinx, 2005).

Finally, there is also research that supports the importance of *the Website Design* features when predicting online purchasing. The design, the possibility to interact with other users and the registration process can help to identify future online purchases (Kim and Kim, 2004). However, variables directly related to the static and interactivity level of the site are not going to be considered in this experiment.

Despite the fact that these 5 main groups are found in the literature to predict online purchases, not all are going to be necessary for the experiment. For instance, whether web design can explain future intention to purchase on-line is not the interest of our experiment. The experiment’s research focus is forecasting, from the customers who are currently purchasing on the physical store, which are the more likely to buy their basket online. Then, historical off-line data and socio demographics will be used.

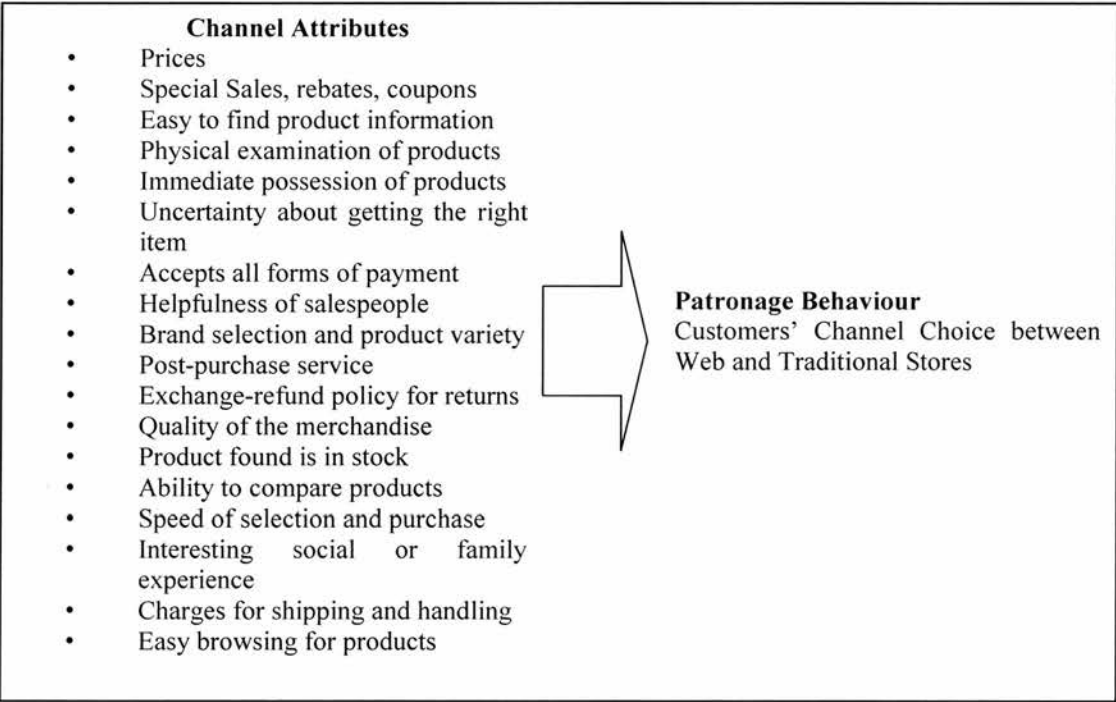
Is it possible to predict and explain customers channel choice between store and Website?

The question is difficult to answer without previous knowledge of the benefits that customers expect from each retail format. To predict and explain the reasons why a specific channel is chosen by a customer or a group of customers, a list of 18 attributes which explain channel choice behaviour is obtained by Chiang, Zhang and

Zhou (2004).

As shown in Figure 6.1, customer patronage behaviour is affected by prices, special sales, etc. In fact, each customer has a preferred combination of attributes. Based on personal ranking, the retail format that best suits the customer’s preferences and needs is more likely to be chosen.

Figure 6.1 List of channel attributes affecting patronage behaviour



Source: Chiang, W-y.K., Zhang, D. and Zhou, L. "Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression." *Decision Support Systems*, October 2004:6.

It is important to mention that there is not a unique preferred combination of attributes by the customer. When attempting to forecast consumer channel choice, type of product and customer’s situation are also two essential aspects.

In reference to the type of product, it is demonstrated that not all the products are suitable for sale on the web (de Kare Silver, 1998). Groceries are one of the most difficult objects of trade for electronic commerce. Not only for the inherent limitation

of perishable goods, but also because the volume of the purchase, the different type of items and the frequency of purchase (Raijas, 2002).

In respect of customer acceptance, the online purchasers are segmented in several ways within the research literature (e.g. Darian, 1987; Greco and Fields, 1991; Lee, Kim and Ahn, 2000; Vellido, Lisboa and Meehan, 1999; Parasuraman and Colby, 2001). According to Parasuraman and Colby (2001), online customers can be split into five basic profiles, according to their internet shopping acceptance. The authors label these as:

‘explorers’,
‘pioneers’,
‘sceptics’,
‘paranoids’ and
‘laggards’.

For instance, explorers are characterised as the first ones to use the Internet; laggards reject the new technology.

Apart from Internet customer acceptance, an alternative segmentation is also proposed by Reibstein (2002). The author suggests finding out which attributes are most important in the consumer choice process, depending on whether it is the initial purchase decision or not.

Based on previous cited research, there is clear evidence that there is the possibility to explain and predict customer channel choice, as long as three main aspects are studied. These are: firstly, customers preferred channel attributes; secondly, customer internet acceptance and thirdly, type of product. Besides this evidence, it is pointed out Burke’s (2002) proposal, as it will be assumed within this entire Chapter. According to Burke (2002), it is not a matter of the channel but the global shopping experience store and website should be considered complementary channels. Then, customers are able to choose between one and another (or both) retail formats,

depending on the type of product/service he/she is interested in buying, the expected benefit of shopping and his/her stage in the purchasing process.

Burke (2002: 426-428) states

'Consumers are not interested in technology for its own sake. People want the basics in their ideal shopping experience (...) Consumers are interested in shopping differently for different types of products. Different shoppers have different needs and wants, and they will go where they are best served. Internet is not a substitute for the physical store but a complement. For instance, the author shows that the main motive to use in-store visit is purchasing and paying for products. Conversely, the main motive to use the Internet is to search for product information.'

Despite the fact that Burke does not specifically distinguish between the on and off-line buying process, publications about the on-line buying process are found in the literature (Häubl and Trifts, 2000; Dholakia and Bagozzi, 2001). The main difference between the two buying processes is based on the power of interactivity, particularly when searching for information provided by online tools. Based on this multichannel retail environment Kaufman-Scarborough and Lindquist (2002) propose a segmentation scheme based on customer's perceptions of convenience. The authors examine five specific non-store channels (the internet, print catalogue, shopping TV, infomercials and advertising) to conclude that some shoppers just want to purchase in the store and reject multiple forms of non-store shopping. However, others prefer to combine more than one channel, although the purchase tends to end in-store. On the contrary, there are pure on-line shoppers as well. Consequently, the channels complementarity concept proposed by Burke (2002) and used in the present study as a theoretical fundamental, is demonstrated by other authors as well.

6.2 Formulating the research problem

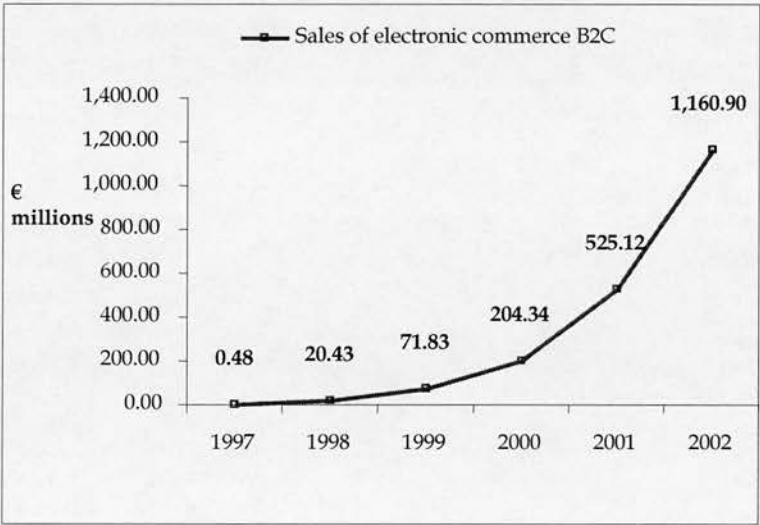
6.2.1 Introduction

The experiment carried out in this Chapter is also based on internal data gathered from the loyalty card and scanner systems of SUPSA.

The appearance of the Internet meant a new challenge for many companies. Particularly in the food retailing sector, it was known that this new technology could generate a considerable change according to the way firms market and distribute their groceries to customers.

Figure 6.2 illustrates that although Spain does not have dramatic figures related to sales over the Internet compared to other countries in Europe, more than €525 millions were achieved from the accumulated electronic commerce to the final customer in Spain in 2001 which was the year to which the first data is referred. (Figure 6.4 illustrates the period of time used for the experiment).

Figure 6.2 Evolution of the Spanish B2C e-commerce sales



Source: AECE-FECEMD (2000)

Although Internet penetration in Spain was relatively slow during 2001, it is still growing. Moreover, there are some Spanish regions where the use of the internet and online purchasing increased considerably. During 2001, Catalonia, Madrid and Vasc Country were the Spanish regions with higher internet penetration, achieving 26.6%, 23.4% and 25.1% respectively (AIMC, 2001). Internet penetration is measured by the number of access, from households, work or university. It is assumed that some people may have Internet connection at work and at home.

This trend of an increase in Internet users and Internet shoppers was taken into account by SUPSA's managers. SUPSA's managers decided to invest in an e-commerce platform.

The decision to launch the e-commerce platform was based on the assumption that current customers would be interested in online grocery shopping. In fact, the process of taking the decision was not trivial. On one hand, there was the fear about what was going to happen with their physical shops if most of the clients decided to buy on-line. With the Internet bubble in 2002 (Chiang, Zhang and Zhou, 2004), it was not strange to think that electronic commerce could lead to cannibalising sales, where current customers shift their business from a real shop to an on-line one.

On the other hand, the Internet was seen as a new alternative for distributing products and services to the customers. In 2000 every day, there was a new e-competitor on the scene. Some competitor's brands were completely unknown, but others were consistent grocery store brands which were expanding their distribution by their website. Although SUPSA was a regional supermarket chain brand, they had leadership in their region and they were interested in keeping their position. Therefore, they decided to launch a Website, even though the penetration of the Internet at that time was really low. The Website's marketing objective was not focused on capturing new customers but not to lose any of the current ones. Therefore, in December 2000, SUPSA launched their electronic supermarket store called 'plusfresh.com' (<http://www.plusfresh.com>).

The company has been investing heavily in developing internal databases, most of them built from actual customer data. The objective was to identify the customers who were going to buy online by using the observed data coming from the loyalty card and scanners systems of the company.

The following sections of the Chapter explain in-depth the research process that was carried out for the experiment.

6.2.2 Setting the research problem

The research problem attempts to answer this main question:

- Which of the current customers are likely to buy online?

6.2.3 Structuring the problem

The main goal of the research problem is to learn from the current customers' specific behaviour to predict their own individual behaviour but in a different situation. Observed historic behavioural data has already proved to be commonly used as an effective predictor (Schmittlein and Peterson, 1994). Particularly in this experiment, observed behavioural data from the physical store is going to be used to forecast online purchasing. As explained in detail in section 6.3.2, data related to shopping convenience are assumed to help in identifying online purchasing. To solve the research problem, 5 steps are required. A full explanation of the implementation approach is provided in section 6.5.2, but a short description of these steps is introduced:

- Learn from in-store past shopping convenience observed behaviour: Once the variables for convenience are established, they are introduced into the software, which starts to learn from them in order to recognising possible patterns of behaviour. It is obvious that the decisive step of selecting the relevant variables is a marketing expert's responsibility.
- Decision of discriminator criteria: Also the criteria which are going to be considered to select the most appropriate pattern of behaviour are decided by the experts parallel to the learning stage.
- Recognising latent patterns of behaviour: As there is not any previous specific known behaviour related to online purchasing, the unsupervised learning algorithm creates many classifications. Each classification presents an undetermined number of segments. Moreover, each segment is likely to contain a

different number of individuals, depending on the classification.

- Identifying online pattern: This is the most crucial stage. The marketing experts are responsible for choosing the classification which they think is the most relevant in terms of providing online purchasing information. Classifications are analysed according the discriminator criteria. Just the ones which fulfil the criteria 100% are selected. Afterwards, the forecasting task is in a strict sense developed.
- Validation of the results: The multichannel theoretical framework recommended by Burke (2002) is followed with this experiment. Accordingly, at this stage, the aim of this research is not only to find out which are the current customers who are likely to buy online, but also to determine who are likely to use both of SUPSA's retail formats. Establishing the adequacy degree of each individual to each segment is possible in this experiment. To evaluate the results, the comparison between LAMDA's forecasts and the actual behaviour will be carried out.

6.3 Obtaining information: Identifying data sources

A specific data base was built for this experiment. The company captured customer's information from the city of Lleida (See Figure 6.3). The city of Lleida was chosen instead of other places where SUPSA has supermarket stores because of town's specific environmental characteristics, which make it the most suitable place to develop the experiment.

Figure 6.3 Location of Lleida



A representative Internet penetration ratio was the first criterion for the selection of the suitable city. Therefore, from all the places SUPSA had store presence, Lleida was the most technological advanced. Council surveys (Mir, Gonzalez and Gil, 2002) revealed that Lleida's Internet ratio was similar to Barcelona's index (26% of the people had Internet access via household, university or work), and this ratio was higher than Catalonia's average rate.

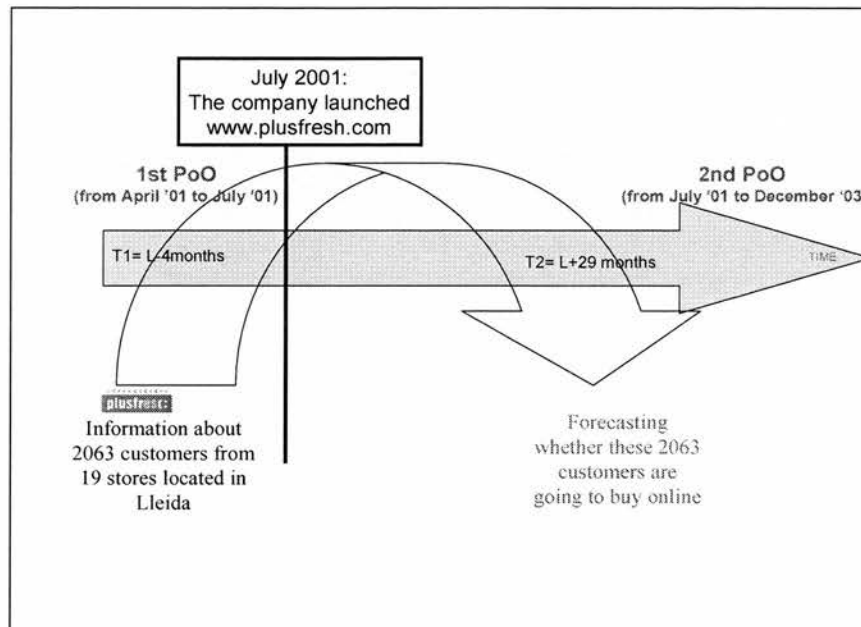
The second criterion was related to socio-demographic and economic features. The importance of socio-demographics indicators has been shown when predicting online buyers. From Spanish Internet surveys (AECE, 2003), the most common Internet user is characterised to be, mostly from urban areas, middle and highly educated, employed, and with medium-high incomes. Lleida is capital of the province, and the population matches the required socio-demographic profile for this second criterion.

6.3.1 Period of observation

As defined in Chapter 5, the period of Observation (PoO) is the moment in time to which the data for the experiment refers. There are two different periods of

observation, as shown in Figure 6.4.

Figure 6.4 Determination of the Periods of Observation (PoO)



The first one (T_1) comprises the four months prior to the launch of Plusfresh.com. The second period is longer. The second period (T_2) is the interval between the moment the company launched the website (July 2001) and the time when there were sufficient customers to test the efficiency of the forecasts. It is important to note that initially, it was thought to take the same period of time (4 months). In fact, a pilot test was developed after waiting 6 months time, but it was felt there was no point in comparing real offline customers with LAMDA's results due to the low Internet rates in Lleida. It is important to mention that from the 23% of the people who had access to the net, just 10% stated to have bought at least once online. And from the online purchases, the domestic grocery shopping was virtually absent (Mir, Gonzalez and Gil, 2002).

Consequently, the decision to increase the period of observation was taken. The presumption was that some present offline customers were likely to become online clients once Lleida's Internet connection ratios increased. Finally, a two and a half time period was established ending in December 2003.

6.3.2 Motivation for online purchasing

A specific database was built for this experiment. 2,063 customers were selected from 19 stores spread across the city of Lleida. In addition, some variables were chosen as predictors for on-line purchasing. The selection of these variables was based on experts' opinion and research publications. To know the experts opinion, two meetings were held. Parallel to the meetings, external secondary data (council surveys, Internet reports) and literature review were also considered.

It is important to note that there was no previous existing data directly corresponding to online purchasing, as it was the first time that the company faced this challenge. Therefore, the objective of the first meeting was to find out which was the main motive that led some customers to buy online instead of visiting the physical stores. A second meeting was then held to explore the availability (within the huge internal database) of existing variables related to the main online shopping motivation. At the end of the first meeting it was agreed *Shopping Convenience* was likely to be the main motivation to buy online, subject to two assumptions being valid.

First assumption: 'Shopping Convenience is not Consumption Convenience.'

Convenience is a fuzzy concept which may take different forms and interpretations. Firstly, consumption convenience is referred to all the products and foods which people normally buy when they don't want or have the time to cook them, such as ready meals from supermarkets or take-aways food for restaurants. In general, consumption convenience takes place when consumers are looking for minimising the efforts and time that they need before and after eating the meal. This type of convenience is not considered in the scope of this research but shopping convenience.

In the literature, shopping convenience is not clearly defined. According to Reimers and Clulow (2000), rather than actually defining the concept of convenience, many researchers simply listed its attributes. To the best of our knowledge, Downs's (1961) was the first research contribution who stated that when seeking convenience,

the shopper sought to minimise three costs: money, time and energy. Furthermore, and approximately 30 years later, Gehrt and Yale (1993) identify the temporal, spatial and efforts dimensions when related to convenience.

Table 6.1 A summary of convenience attributes from a literature review

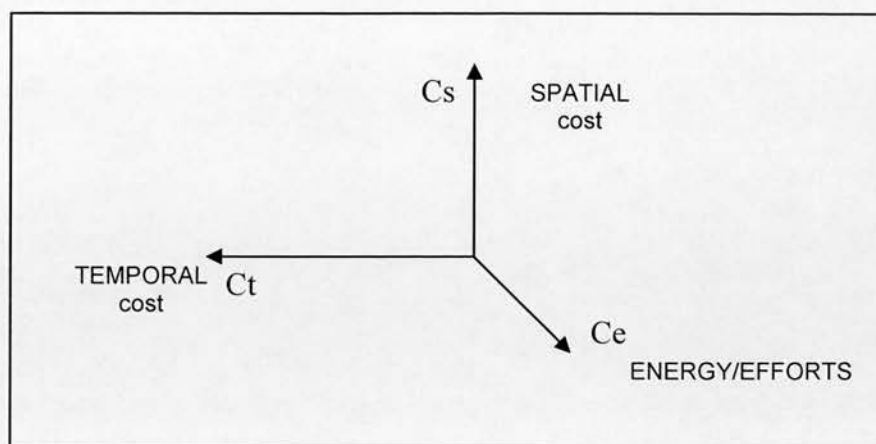
Research Studies	Attributes of convenience						
	Trading Hours	Proximity	Travel time/Access	Internal layout	Parking	Enclosure	Merchandise Variety
Bellenger, Robertson and Greenberg (1997)		*					
Spencer (1978)		*					*
Howell and Rogers (1980)		*	*				
Cymrot, Gelber and Cole (1982)			*		*		
Timmermans, Van der Heidjen and Westerveld (1982)		*		*	*	*	
Bucklin and Gautschi (1983)				*			
Oppewal, Louvieve and Timmermans (1994)			*		*		
Berrell (1995)							*
Kaufman-Scarborough (1996)							*
Bell (1999)	*	*		*	*		

Source: Reimers, V. and Clulow, V. Shopping and Convenience: A model for retail centres. ANZMAC Conference, 2000:8.

As shown in Table 6.1, seven different categories of convenience in terms of trading hours, proximity, travel time and access, internal layout, parking, enclosure and merchandise variety have been studied during the last 30 years. All attributes refer to temporal, spatial or energy dimensions, introduced by Gehrt and Yale. However, Timmermans, Van der Heidjen and Westerveld (1982) and Bell (1999) are the only two that take the global three dimensional meaning of convenience into account.

On the other hand, as far as the table shows, proximity has been the aspect most directly related to convenience. It is important to note that mentioned publications about shopping convenience were mainly focused on off line retail shopping. However, in this study, the three dimensions of convenience introduced by Gehrt and Yale (1993) are followed as our theoretical approach.

Figure 6.5 Schematic concept of Shopping Convenience



As shown in (Figure 6.5), the closer location to the centre, the greater is the shopping convenience and the lower is the total cost ($C_s + C_t + C_e$). A critical part of the experiment is to find out the variables which inherently refer to at least one of shopping convenience's dimensions.

It is important to mention that customers may look for convenience of purchasing, which does not necessarily mean that they also look for convenience of consumption. And the other way round.

Second Assumption: 'Shopping Convenience is the main motivation when alternative channels are offered by one firm'.

To explain this assumption, reference is made to the list of 18 attributes (See previous Figure 6.1). The fact that both the physical store and online store offer exactly the same products and services neutralises some of the possible customer channel preferences. For instance, prices, special rates and coupons, brand selection and product variety are attributes that remain the same, in the physical store and the Website. Furthermore, possible fears relating to uncertainty about getting the right item, exchange – refund policy for returns, quality of the merchandise, product found in stock, also disappear, as the current SUPSA's customers know the company. Therefore, these attributes also remain the same in both channels. The rest of the attributes from the list are directly related to shopping convenience. Some of them

seem to be more relevant when buying in the traditional store, such as physical examination of products, helpfulness of sales people. Conversely, others are certainly related to online purchasing such as easy browsing for products and ability to compare products.

Consequently, it is going to be assumed that the main motivation (in a case when the grocery firm is the same for both channels) that can make the current customer change from SUPSA's traditional outlet to SUPSA's online supermarket is shopping convenience. The higher value the customer assigns to shopping convenience (instead of other shopping benefits), the most potential exists to buy online.

In the research literature, publications aiming to compare online shopping and store shopping were also considered when reaffirming the second assumption. Despite the fact that several analyses of the advantages and disadvantages of both retail formats (also called channels) are captured (Strader and Shaw, 1997; Breitenback and Van Doren, 1998; Crawford, 2000; Degeratu, Rangaswamy and Wu, 2000; Ray, 2001; Burke, 2002) most of the publications directly or indirectly conclude with a common denominator: Although being defined in different ways, convenience is also higher ranked for online shopping than store shopping (Kalakota and Whinston, 1997; Burke, 2002; Dahlén and Lange, 2002).

According to Raijas (2002), convenience is expected for either physical store customers or store website customers. However, the website not only can achieve the traditionally most important factors affecting the physical store choice (low price level, customer service, location, product assortment) but also it is able to avoiding all the inconvenience of grocery shopping (looking for the products, self picking, waiting queues, self delivering .etc). Based on an electronic grocery shoppers survey, Raijas (2002:111) concludes that

'the principal benefits for online purchasing are (1) time and effort saving, (2) time and place independence and (3) possible tools for follow-up and planning'.

External secondary sources supported the assumptions. For example, according to AECE's (2000; 2003) results, the first Internet users' motivation to buy online is convenience.

When the main online shopping motivation was finally determined, the second meeting selected, from the huge internal company databases, the variables that could define this accepted motive.

At this point, it is important to note that the information stored in the internal database was mainly behavioural and socio- demographic. There was no data corresponding to either customers' Internet perceptions or customers' online purchasing intentions, as data all come from the loyalty card programme and scanner systems.

6.3.3 Predictors for '*Shopping Convenience*' used in this study

In order to identify shopping convenience, 28 indicators were initially selected for the experiment. Noting that by merging different categories of data, the predictive power of the modelling exercise is maximised (Montgomery, 2001); the predictors were split into two main categories. The first was termed 'socio-demographic details'. The second one grouped the observed 'behavioural in-store data'. It is important to note that all these indicators correspond to the previous fourth months to the launch of Plusfresh.com (Period of Observation T1).

Socio-demographic details

As previously mentioned, customer demographics have been extensively applied to explain and forecast online purchasing. Consequently, 6 demographic predictors available in the internal database were originally selected. The 6 variables include both the information related to the cardholder subscriber and his/her household. As it is listed in Table 6.2, the variables mainly focused on the individual characteristics are *customer code*, *age*, *gender*, *employment status*. Mention that *customer code* variable is used to identify each customer (or household). Then, although it is not

used for the learning process itself, it is the key data for comparing and analysing the results in the final stage.

Directly related to the household are V5 (address) and V6 (email). There was no information available related to the household income. Then, the address was chosen as it was likely to provide information about the area where the customer lives (rich, medium, poor area of the city). Variable 6 was also interesting for the purpose of the experiment. Having *an email* was assumed to be positively related to having Internet connection and frequent accessibility. Hence, it seemed to be a highly significant variable.

Table 6.2 Socio demographic predictor variables selected for the experiment

N	Variable name	Description
Information related to the individual Cardholder		
1	Customer code	The loyalty card code which allows to identify the customer
2	Employment Status	There is a classification between: housewife, retired, unemployed, employed, employee, self-employed
3	Gender	Women/Men
4	Age	Date of Birth
5	Address	This variable gives information about the area where the customer lives,.
6	Email	Yes or No. The answer indicates whether he/she has Internet access and frequency of access.

Despite the fact that race and language have been also considered interesting predictors (Padmanabhan, Zheng and Kimbrough, 2001), these were not applicable in Lleida, where the major spoken languages are Catalan and Spanish. Moreover, despite the fact that immigration is slightly increasing, SUPSA’s customers are dominantly Spanish.

Observed behavioural in-store aspects

As is listed in Table 6.3, the *Observed buying behaviour data* captured from purchases in traditional stores joins 21 predictors classified according to the 3 shopping convenience dimensions: *Temporal*, *Energy/effort* and *Spatial* categories. The objective of this selection is to establish a concrete description of the shopping convenience concept for the experiment.

Despite the fact that almost all the indicators could be located in more than one category, possible overlapping is ignored. Accordingly, each predictor has been only located to one single dimension.

Table 6.3 Classification of the observed behavioural in-store aspects.

DIMENSIONS OF CONVENIENCE (Gerht and Yale, 1993)	Number of predictors
TEMPORAL DIMENSION	5
ENERGY EFFORT DIMENSION	9
SPATIAL DIMENSION	7
TOTAL CONVENIENCE CONCEPT	21

As shown in Table 6.4, the *temporal dimension* is represented by 5 predictors. When individuals experience high levels of time scarcity, they are likely to have certain ways of thinking about and using time (Kaufman-Scarborough and Lindquist, 2003). Therefore, a selection of specific indicators able to deduce degrees of time scarcity has been chosen for the experiment such as % delivered purchases after seven p.m. Indicators more related to frequency are collected as well. According to Raijas (2002), online shopping frequency is lower than in a conventional grocery store. Based on his statements, variables 7, 8, 9, 10 and 11 were selected.

Table 6.4 Observed behavioural in-store data predictors selected for the experiment: Selection of the convenience’s temporal dimension predictors

N	Variable name	Description
TEMPORAL DIMENSION		
7	% delivered purchases after 7 p.m.	Percentage of monetary value referred to the deliveries after 7 pm (according to the total amount spend during T1)
8	Mean No. of days per week on which purchases are made	Average number of trips to the shop by week
9	% purchases made on Saturdays (number of customer trips)	Percentage of monetary value made on Saturday (from the total shopping trips made during T1)
10	% purchases made on Saturdays (amount)	Percentage of monetary value spend on Saturdays (from the Total Purchase of T1)
11	% of purchases made from Monday to Wednesday (amount)	Percentage of monetary value spend from Monday to Wednesday (from the Total Purchase of T1)

In reference to *effort/energy dimension*, 9 predictors were determined (See Table 6.5). According to Raijas’s (2002) contributions, despite the fact that online grocery

shoppers tend to buy the same products as in a conventional store, they tend to concentrate purchases of dry products and beverages. On the one hand, *fresh products* are bought less online. On the other hand, avoiding the picking and handling were also relevant points when choosing online grocery shopping. Supporting this, variables from 12 to 19 were selected. The number of outlets where the customer purchase also informs about an additional effort, therefore whether the customer buys in more than one is captured in this variable (V20).

Table 6.5 Observed behavioural in-store data predictors selected for the experiment: Selection of the convenience's energy/efforts dimension predictors

ENERGY EFFORTS DIMENSION		
12	% purchase of fresh produce /Total purchases	Percentage of monetary value spend fresh produce from the total items purchased
13	% meat purchases at self-service counter / Total purchases	Percentage of monetary value spend on meat at self service from the total items purchased
14	% of purchases made up by special offers/ Total packed product	Percentage of monetary value spend on special offers from the total packed items purchased
15	% delivered purchases (amount)	Percentage of monetary value delivered at home from the total amount spend on T1
16	% delivered purchases (n° of customers trips)	Percentage of delivered purchases at home from number of customer trips to the shop during T1
17	Was auto-scanning used?	Some SUPSA's stores have auto-scanning service. This data informs whether it was used by the customer or not.
18	Means of transport	Based on the address information this variable indicates the distance to a store. This variable informs whether the customer comes by food, walking or by car.
19	% of coupons and discounts redemption	From all the coupons and discounts launched by the company, this variable measure the % that the customer uses them.
20	Number of outlets where customers shop	Number of SUPSA's stores used by each customer

Particularly, an explanation of *was scanning used* (V17) and *meat purchases at self service* (V13) is required as they are special features of the company. Auto-scanning is not available in every store. Customers are given an easy-to-use device which helps them to save time when checking out. Customers do not need to wait to have their items scanned, because they have already scanned their items while they were walking and picking them from the aisle. Information corresponding to *Meat purchase at self service* is captured by the scanner systems every time the customer takes it directly to the meat shelf instead of waiting for his turn in the Butchery inside the supermarket store.

Special sales and coupons was one of the most relevant attributes of performance at web stores (Chiang, Zhang and Zhou, 2004). Therefore, variable 14 (*% of purchases made up by special offers and % of coupons and discounts redemption*) were also included for defining the effort dimension of convenience. *Means of transport* (V18) was resulted from the transformation of the variable *address*. This resulted variable was split between 3 categories, which include ‘by foot’, ‘walking’ and ‘by car’. Section 6.4.1 explains the transformation process that was carried out.

Referring to the *spatial dimension* of convenience, Table 6.6 shows the variables that correspond to the location and distribution of the products in the store. In particular, 4 indicators were selected to describe this type of convenience category (V24, V25, V26 and V27).

Table 6.6 Observed behavioural in-store data predictors selected for the experiment: Selection of the convenience’s spatial dimension predictors

SPATIAL DIMENSION		
21	Average spent per item (Total T1)	The money spent by item purchased during T1 divided by the number of items
22	Average purchase (per total trips)	From all the purchases that the customer realises during this period, this variable shows the average of purchases
23	Total Purchases	Total monetary amount of spending during T1
24	Size of the outlet	SUPSA’s classify their stores into big, medium or small.
25	No. of different articles purchased in the period	Number of different items purchased during T1
26	No. of departments where no purchases were made	Number of departments where NO item was bought during T1
27	Number of TOTAL items purchased	Number of references bought during T1

Furthermore, Raijas (2002) suggested that the average amount spent in electronic grocery shopping was generally higher than the amount spent in the store. Based on this suggestion, *Average of items purchased*, *Average purchase with the company* and *Total Purchases* were taken into account for the experiment and included in the spatial dimension (See Table 6.6).

Predicted variable: ‘Online purchasing behaviour’

No previous SUPSA based historical data corresponding to online purchasing existed, so an assumption was required before the implementation of LAMDA’s

approach.

- Assumption: The customers who previously bought by distance selling (fax, telephone, email) are considered the strong potential Internet buyers.

SUPSA had information on the customers who had ordered their purchases either by fax, email or telephone. 78 out of 2.063 had demonstrated attributes that showed them to be more interested in shopping convenience than in other store benefits. These 78 customers were noted and labelled as distance-buying individuals.

Although a deeper explanation is provided in section 6.5.2, we would like to remark the important role of V28 (See Table 6.7). This variable is not considered in the learning process. It is only going to be used as a criterion to decide the most relevant classification from the wide range of results provided by the unsupervised LAMDA algorithm. V28 is the predictive variable, crucial when forecasting the online customers.

Table 6.7 Determination of the predictive variable

N	Variable name	Description
Predictive Variable		
28	Has the customer ever purchased by distance selling?	Despite the fact that the company did not have a Website, their customers had been able to order their shopping basket by telephone, fax or email.

6.3.4 Validating corpus

The variables presented in this section correspond to the second period of observation (T2). In December 2003, the company checked the customers who bought, at least once at www.plusfresh.com (See Table 6.8).

Table 6.8 Variables related to the real online purchasing customer

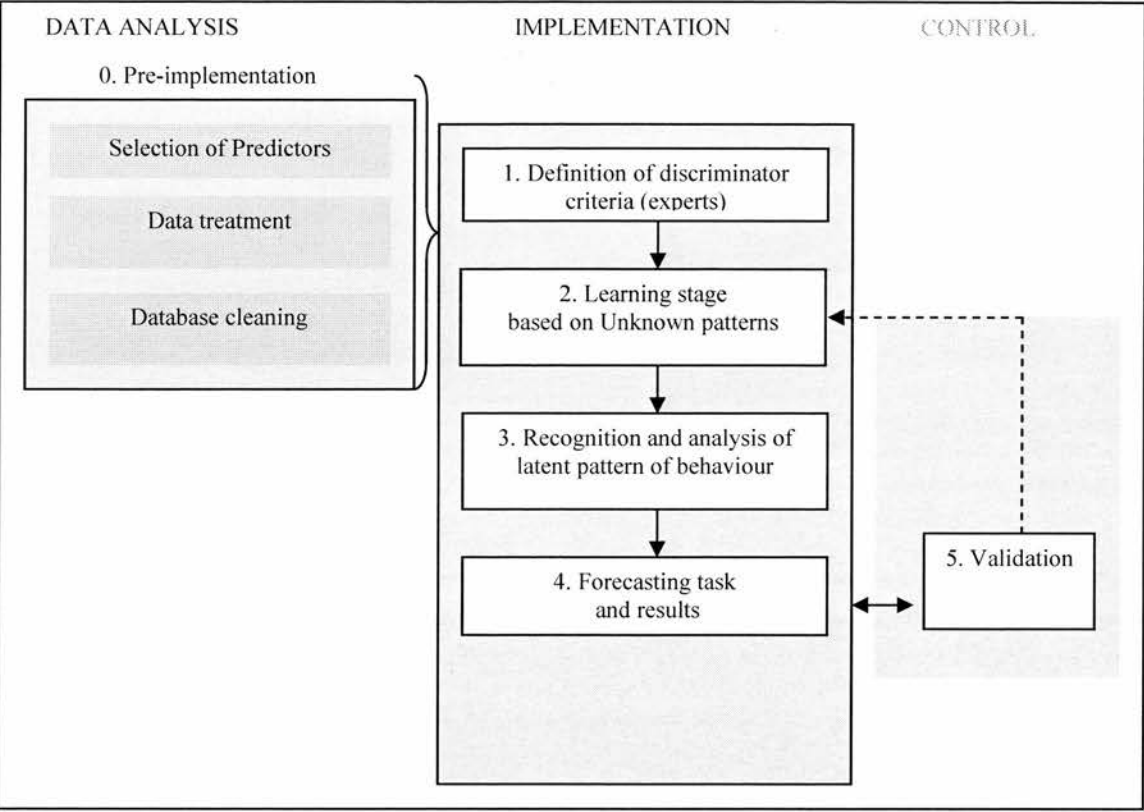
N	Variable name	Description
Real online behaviours (Variables related to PoO T2)		
29	Has the customer ever bought at www.plusfresh.com?	Yes or No (from July 2001 to December 2003)

V29 informs whether the customer has bought at least once at plusfresh.com since it was launched in July 2001.

6.4 Implementation of the unsupervised forecasting model

As with the supervised forecasting model, LAMDA’s unsupervised forecasting model also follows 3 main phases. As illustrated in Figure 6.6, the first phase is the data analysis. Data analysis is required prior to the implementation. Within this phase, similarly to the data analysis of the supervised forecasting model, the final predictors are chosen, the type of data is classified, the proper transformation of data considered and data cleaning carried out. Afterwards the implementation itself is developed. The 4 sub-stages which compose the implementation phase in the unsupervised forecasting model include the definition of the ‘discriminator criteria’ (See section 6.4.2), the learning stage, the analysis of the recognised latent patterns of behaviour and forecasting task’s results. Once the results are known, the final validation phase is then developed as the control phase.

Figure 6.6 LAMDA’s unsupervised forecasting model



6.4.1 Data Analysis

Before implementing the unsupervised learning model, a specific analysis of data is developed to reconsider the initial 28 selected predictors. Similar to Experiment 1, once the final predictors are selected the data treatment is carried out. This phase also encloses the database proper cleaning.

As far as the selection of final predictors is concerned, the decision to ignore the predictors related to the individual cardholder was taken. Mention of the decision of moving the cardholders’ variables apart was mainly based on experts’ judgments. Despite the many publications within the literature which suggested that indicators such as age, gender and employment status were powerful when forecasting online purchases, they were rejected for this research. There were two main reasons to exclude these variables. The first one was based on the fact that the individual was not the subject of the research. His/her household was. The second reason and related to the first one is that the most common Internet user profile described in section 6.3

is likely to be helpful when analysing the individuals who live alone. In section 6.3, the most common Internet user profile was characterised to be mostly from urban areas, middle and highly educated, employed and with medium high incomes. However, when analysing households, these variables are not as much interesting as any member of the house can share one or more of these attributes.

Related to data treatment, mention that the *address* variable was used to determine a new variable, *means of transport*. In fact, the specific address alone did not provide any interesting information. However, when using it to find out the distance between each customer's home and the closest store, this information became essential. Once this information was known, each distance was turned into 3 possible categories, 'by foot', 'walking' and 'by car'. When there was a high proximity to the store, it was assumed that the customer would take less than 5 minutes to go to the store. Then, he/she was assigned 'by foot'. When the customer needed between 5 to 10 minutes to arrive to the store, it was assumed that he/she came 'walking'. For higher distances, the customer was assumed to come to the store by car. Noting at this point that a variable that initially was qualitative (address) was firstly transformed in quantitative (distance), and then discretised to a order of magnitude variable (means of transport).

In addition, LAMDA can deal with nominal and numerical types of variables simultaneously. Table 6.8 illustrates the final predictors, according to the qualitative or quantitative variables.

Table 6.9 Qualitative and quantitative data

QUALITATIVE Predictor	QUANTITATIVE Predictor
Email	% delivered purchases after 7 p.m.
Size of the outlet (OM)	Mean No. of days per week on which purchases are made
Was auto scanning used?	% purchases made on Saturdays (number of customer trips)
Means of transport (OM)	% purchases made on Saturdays (amount)
Has he ever bought using distance selling options?	% of purchases made from Monday to Wednesday (amount)
	% purchase of fresh produce /Total purchases
	% meat purchases at self-service counter / Total purchases
	% of purchases made up by special offers / total packed product
	% of coupons and discounts redemption
	% delivered purchases (amount)
	% delivered purchases (n° of customers trips)
	Number of TOTAL items purchased
	Number of outlets where customers shop
	Average spent per item (Total T1)
	Average purchase (per total trip)
	Total Purchases
	No. of different articles purchased in the period
	No. of departments where no purchases were made
5 qualitative (2 are OM)	18 quantitative

The relationship between 5 qualitative predictors and 18 quantitative was suitable for the research. Various pilot tests were developed to test whether a higher number of qualitative variables were needed. However, pilot tests results showed that there was no significant difference when using a higher number of nominal variables. Consequently, as carried out with *address* variable, data treatment stage was just oriented to turn some data into more precise information.

Related to database cleaning it was practically not relevant in this experiment as all the initial customers 2.063 was used for the research. There were no missing values.

6.4.2 Implementation of unsupervised forecasting model

Once the data analysis phase is finished, the implementation of unsupervised forecasting model starts. Due to the large number of classifications which will result

from the learning and recognition stages, the first step is focused on defining the discriminator criteria which will help to choose between them. The learning task is performed in the second stage. Afterwards, the recognition of latent patterns of behaviour and its analysis is developed. The fourth step consisted of the forecasting task. Finally the results were collected.

Defining the discriminator criteria

Unlike the supervised model, measuring and labelling the existing patterns of behaviour was not required as no previous pattern of behaviour existed. As Figure 6.6 illustrates, the unsupervised learning algorithm first learns from customers' information, recognises afterwards latent patterns of behaviour and suggests a set of possible classifications to be analysed. In order to select between the vast options provided by the model, some criteria to discriminate between the possible classifications has to be decided at this point. A meeting with marketing personnel from SUPSA was held. The purpose of the meeting was to know the decision process and the criteria that experts would use to identify and forecast 'online shopping behaviours' without disposing of any particular market research survey. Interesting conclusions resulted from the meeting and 3 discriminator criteria were established. The selection of the most appropriate classification was carried out by means of these next criteria:

1. The classification has to follow the grouping rule

The grouping rule is measured by the concentration of the distance-selling individuals in one or more segments of the classification. Experts agreed with the fact that the customers who had purchased by distance selling would be used to determine the customers who would buy online. Then, the classification that presented at least one segment with a high concentration of distance buyers would be considered. Particularly, if one of the segments of the classification joined more than 25% of the -buying individuals (that means more than 21 distance buyers in absolute numbers), the classification was marked as 'interesting'.

2. The classification has to be manageable

Apart from the grouping rule, the classification should be manageable. The unsupervised learning algorithm suggests a wide set of classifications based on the latent pattern of behaviour that the algorithm recognised. Then, each classification is likely to be composed by different numbers of segments. From management experts' point of view, it was decided that a classification which showed more than 5 segments was considered not manageable as it is difficult to be interpreted.

3. The classification has to be balanced

Classifications resulted from the recognition stage are also likely to present unbalanced segments. A classification is unbalanced when one of its segments groups has more than the 80% of the total individuals (2.063). In that case, there is virtually just one segment, and when interpreted from a marketing point of view it is not useful as little discrimination between individuals is provided.

In addition to these 3 criteria, there are also 2 conditions inherently related to LAMDA which have to be considered as well. These are the following:

4. The classification has to be stable

A stable classification takes place if at the recognition stage, all individuals are reassigned in the same segment. When applying the unsupervised learning, the individuals who are located firstly to some segments have a higher weight to the segment than the last ones. Once the algorithm learns, the individuals tend to be located in the same segment. When the learning process assigns repeatedly the same individuals in the same segments, the classification is stable.

5. The classification has to be unique

The classification is likely to be proposed more than once. Different combination

between LAMDA’s capabilities may suggest exact patterns of behaviour. All the repetitive classifications are directly removed by this criterion.

Having explained the discriminator criteria, it is important to note that just the classifications which fulfil the five criteria will be selected as promising.

Learning stage

The unsupervised learning process takes place when the different types of hybrid connectives provided by LAMDA are combined automatically with a specific level of tolerance. This is a trial and error approach based on an iterative process. Then, several combinations of three elements which include a specific number of iterations, an applied fuzzy connective and an established level of tolerance are tested. The learning process is considered to be finished when either the classifications are stable or a pre-determined number of iterations are carried out. In this experiment, 10 is the pre-determined number of iterations.

Analysis of Recognised latent pattern of behaviour

Using the unsupervised learning capabilities of LAMDA algorithm, 945 classifications were obtained. Table 6.10 shows the combinations applied for the experiment. Based on the same number of iterations and with an automated tolerance selected, minmax algorithm was able to recognise 815 different latent patterns of behaviours while probabilistic algorithm just recognised 7.

Table 6.10 Classifications resulted from LAMDA’s unsupervised learning capabilities

Fuzzy connective	Number of iterations	Tolerance	Number of classifications resulted
MINMAX	10	Automatic	815
FRANK	10	Automatic	123
PROBABILISTIC	10	Automatic	7
TOTAL			945

The number of classifications resulted from each algorithm is not important. What really matters is how many classifications accomplish the discriminator conditions. The following table explains the finalist classifications according to the unique and stable criterion:

Table 6.11 Finalist classifications based on unique and stable criteria

		UNIQUE CRITERION	STABLE CRITERION
Fuzzy connective ¹	Number of unique classifications obtained	Number of unique classifications	N° of Stable classifications from the total obtained
MINMAX	815	608	29
FRANK	123	98	3
PROBABILISTIC	7	7	1
TOTAL	945	713	33

Despite the fact that initially, 945 classifications were obtained, just 713 were not repetitive classifications. Noting the initial 815 classifications suggested by minmax algorithm that were reduced to 608. Moreover, when analysing the stability of these 713 classifications, a dramatic reduction of the number of possible finalists took place. Just 33 classifications out of 713 accomplished the stable criteria as well, 29 came from Minmax, 3 from Frank and just one from Probabilistic. The next step was to analyse the 33 classifications according to the rest of the criteria. As it is shown in Table 6.12, all the classifications recognised by FRANK and PROBABILISITIC presented less than 5 segments. However, just 18 classifications out of the 29 resulted from MINMAX satisfied this managerial criterion.

Table 6.12 Finalist classifications based on the managerial criterion

N° SEGMENTS	2	3	4	5	6	7	+ 7	FULFILED MANAGEABLE CRITERION
MINMAX	7	0	2	9	2	3	6	18
FRANK	1	1	0	1	0	0	0	3
PROBABILISTIC	1	0	0	0	0	0	0	1
TOTAL								22

¹ The number of unique and stable classifications obtained by Lukasiewicz was 0.

Then, the balance criterion was analysed as well. Firstly, the 18 classifications from minmax were analysed. None satisfied the balanced criterion. Similarly with the probabilistic approach. As far as the 3 FRANK classifications are concerned, one was removed as more than 80% of the total individuals were grouped in the same segment. Table 6.13 shows the only two classifications that fulfilled all the criteria.

Table 6.13 Finalists classifications based on balanced criteria

		BALANCED CRITERION					FULFILED CRITERIA
Classification	Tolerance	% of individuals grouped in each segment					The 4 th first criterion are fulfilled
2. FRANK	0.443	54%	21%	17%	14%	3%	YES
3. FRANK	0.454	66%	30%	4%			YES

Finally, the last criterion (grouping criterion) was checked in order to know whether the classification was suitable to forecast online purchases. Both, FRANK 0.443 and FRANK 0.454 were analysed:

Table 6.14 Composition of classification 2, Frank 0.443

2. FRANK 0.443			Total segment	% distance buyers
Segment 1	Distance buyers	22	928	28%
	Others	906		
Segment 2	Distance buyers	20	427	26%
	Others	407		
Segment 3	Distance buyers	8	357	10%
	Others	349		
Segment 4	Distance buyers	25	285	32%
	Others	260		
Segment 5	Distance buyers	3	66	4%
	Others	63		
TOTAL	Distance buyers	78	2063	100%
	Others	1985		

As illustrated in Table 6.14, segment 1, segment 2 and segment 3 present a concentration ratio of distance buyers of 28%, 26% and 32% correspondingly. The grouping principle is perfectly fulfilled.

Table 6.15 Composition of classification 3, Frank 0.454

3. FRANK 0.454			Total segment	% distance buyers
Segment 1	Distance buyers	38	1353	49%
	Others	1315		
Segment 2	Distance buyers	36	625	46%
	Others	589		
Segment 3	Distance buyers	4	85	5%
	Others	81		
TOTAL	Distance buyers	78	2063	100%
	Others	1985		

Also segment 1 and segment 2 from classification 3 (See Table 6.15) accomplish the grouping criterion as both of them concentrate more than 25% of the total distance buyers.

Forecasting task

It is important to remark that all the individuals placed in the same segment shares the same pattern of behaviour, although this pattern is unknown. For the forecasting task, variable V28 was mainly used. Since the experiment had no ideal partition to conduct a comparison, the 78 clients who had engaged in a remote buying act by e-mail or fax, called distance-buying customers were assumed to be almost certain customers for web purchase. Accordingly, customers who were located in the same segment as the distance buyers were considered potential online customers as well. Based on that, from classification 2 (See Table 6.13), segments 1, 2 and 4 would be selected as the online buyers. That means that the 928, 427 and 285 customers respectively would be considered potential online buyers. From classification 3 (See Table 6.14), the customers located in segment 1 and 2, which are 1353 and 625 respectively would also be the potential online buyers.

From the two final classifications, a selection of just one was required. However, there were no objective grounds for making a final selection between the two because they did not have the same number of segments. To solve this apparent tie-break, we tried to reduce the 5 segments from Classification 2 to 3 segments so a final comparison was then enabled. The reduction of the number of segments is

carried out by the intervention process.

In LAMDA, each individual belongs to all the segments to a greater or lesser extent. Although each individual is finally allocated to the segment to which it presents the greater maximum GAD, it still belongs to the rest of the segments with a specific GAD to each segment. Therefore, each individual is thus assigned an adequacy rating (vector). Looking at the individual vector, it is possible to observe what happens if we multiply one of these adequacy ratings by a corrective parameter. Table 6.16 shows the corrective parameter and how it causes the vector to change, forcing the individual into another segment.

Table 6.16 Intervention process applied in classification 2, Frank 0. 443

Segment 1	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	875	928	1053	1176	1226	1249	1288	1347	1426	1518	1595
2	465	423	356	304	285	282	276	262	246	226	208
3	362	357	348	313	310	293	276	255	220	184	151
4	295	289	242	207	181	178	166	143	118	86	63
5	66	66	64	63	61	61	57	56	53	49	46
Segment 2	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	975	928	882	862	857	848	836	812	764	699	622
2	315	423	517	620	657	685	728	793	893	1032	1147
3	358	357	353	315	310	293	278	259	226	193	176
4	348	289	248	204	178	176	164	143	127	90	72
5	67	66	63	62	61	61	57	56	53	49	46
Segment 3	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	928	928	923	912	896	846	762	653	569	469	356
2	425	423	421	383	363	347	330	295	265	231	181
3	310	357	400	484	536	605	731	904	1037	1218	1406
4	334	289	254	219	204	201	181	153	137	96	74
5	66	66	65	65	64	64	59	58	55	49	46
Segment 4	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	954	928	915	890	854	832	792	724	633	560	497
2	442	423	349	293	275	260	229	207	182	140	90
3	375	357	310	299	270	226	185	162	146	124	94
4	225	289	425	519	606	691	807	926	1063	1207	1357
5	67	66	64	62	58	54	50	44	39	32	25
Segment 5	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	928	928	927	922	917	908	896	871	821	754	674
2	425	423	418	384	365	361	352	339	322	299	277
3	357	357	353	316	310	293	278	259	226	194	176
4	290	289	268	230	205	203	189	167	147	108	84
5	63	66	97	211	266	298	348	427	547	708	852

For instance, when multiplying all the 2.063 GADs in segment 1 for an increasing corrector parameter, it is shown that the customers who initially belonged to segment 2 and segment 4 moved to segment 1 more quickly than the rest. For instance, when multiplying the GADs in segment 1 for a 1.02 corrector parameter, the 423 customers in segment 2 are reduced to 356 and the 289 customers in segment 4 are reduced to 242. These 67 and 47 customers respectively moved to segment 1. Table 6.17 shows the increasing or decreasing tendency of each segment in absorbing or losing customers. Each cell shows the accumulated percentage resulting from applying the corrector parameter.

Table 6.17 Accumulated percentages of the absorbed or lost customers per segment

Segment 1	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	875	928	0.1347	0.2672	0.3211	0.3459	0.3879	0.4515	0.5366	0.6358	0.7188
2	465	423	-0.158	-0.281	-0.326	-0.333	-0.348	-0.381	-0.418	-0.466	-0.508
3	362	357	-0.025	-0.123	-0.132	-0.179	-0.227	-0.286	-0.384	-0.485	-0.577
4	295	289	-0.163	-0.284	-0.374	-0.384	-0.426	-0.505	-0.592	-0.702	-0.782
5	66	66	-0.03	-0.045	-0.076	-0.076	-0.136	-0.152	-0.197	-0.258	-0.303
Segment 2	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	975	928	-0.05	-0.071	-0.077	-0.086	-0.099	-0.125	-0.177	-0.247	-0.33
2	315	423	0.2222	0.4657	0.5532	0.6194	0.721	0.8747	1.1111	1.4397	1.7116
3	358	357	-0.011	-0.118	-0.132	-0.179	-0.221	-0.275	-0.367	-0.459	-0.507
4	348	289	-0.142	-0.294	-0.384	-0.391	-0.433	-0.505	-0.561	-0.689	-0.751
5	67	66	-0.045	-0.061	-0.076	-0.076	-0.136	-0.152	-0.197	-0.258	-0.303
Segment 3	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	928	928	-0.005	-0.017	-0.034	-0.088	-0.179	-0.296	-0.387	-0.495	-0.616
2	425	423	-0.005	-0.095	-0.142	-0.18	-0.22	-0.303	-0.374	-0.454	-0.572
3	310	357	0.1204	0.3557	0.5014	0.6947	1.0476	1.5322	1.9048	2.4118	2.9384
4	334	289	-0.121	-0.242	-0.294	-0.304	-0.374	-0.471	-0.526	-0.668	-0.744
5	66	66	-0.015	-0.015	-0.03	-0.03	-0.106	-0.121	-0.167	-0.258	-0.303
Segment 4	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	954	928	-0.014	-0.041	-0.08	-0.103	-0.147	-0.22	-0.318	-0.397	-0.464
2	442	423	-0.175	-0.307	-0.35	-0.385	-0.459	-0.511	-0.57	-0.669	-0.787
3	375	357	-0.132	-0.162	-0.244	-0.367	-0.482	-0.546	-0.591	-0.653	-0.737
4	225	289	0.4706	0.7958	1.0969	1.391	1.7924	2.2042	2.6782	3.1765	3.6955
5	67	66	-0.03	-0.061	-0.121	-0.182	-0.242	-0.333	-0.409	-0.515	-0.621
Segment 5	0.98	1	1.02	1.04	1.06	1.08	1.1	1.12	1.14	1.16	1.18
1	928	928	-0.001	-0.006	-0.012	-0.022	-0.034	-0.061	-0.115	-0.188	-0.274
2	425	423	-0.012	-0.092	-0.137	-0.147	-0.168	-0.199	-0.239	-0.293	-0.345
3	357	357	-0.011	-0.115	-0.132	-0.179	-0.221	-0.275	-0.367	-0.457	-0.507
4	290	289	-0.073	-0.204	-0.291	-0.298	-0.346	-0.422	-0.491	-0.626	-0.709
5	63	66	0.4697	2.197	3.0303	3.5152	4.2727	5.4697	7.2879	9.7273	11.909

The intervention process is applied to each segment from classification 2. The objective is to observe whether individuals tend to move to other segments when a corrector factor is applied in each segment. Graphically, Figure 6.7 illustrates the intervention process applied in segment 1. As mentioned, all the individuals have a GAD in segment 1. However, just the customers with the maximum GAD to segment 1 are located in it. When all the 2063 GADs in segment 1 are multiplied by the corrector parameter, the individuals which initially had a maximum GAD in segment 2 and segment 4 tends to change to segment 1.

Figure 6.7 Results from the application of intervention process in segment 1

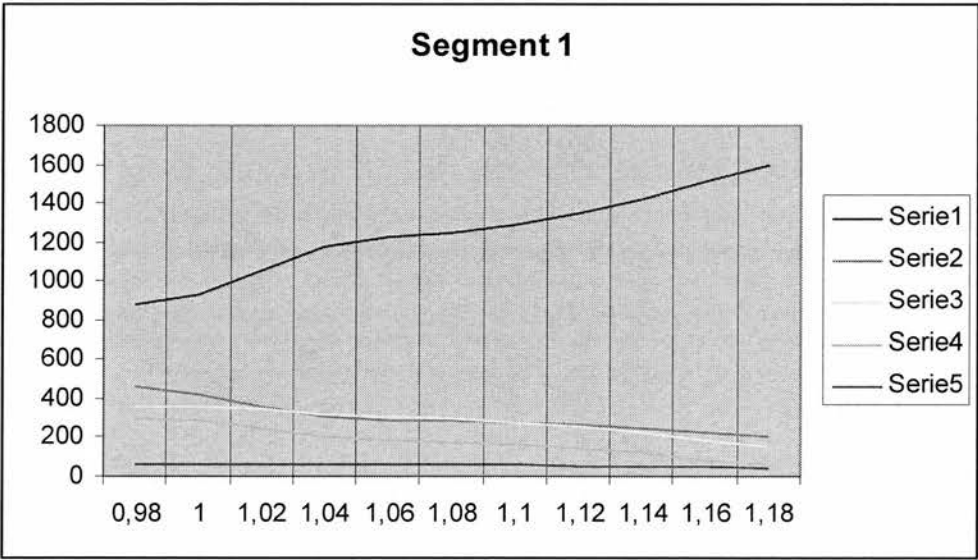
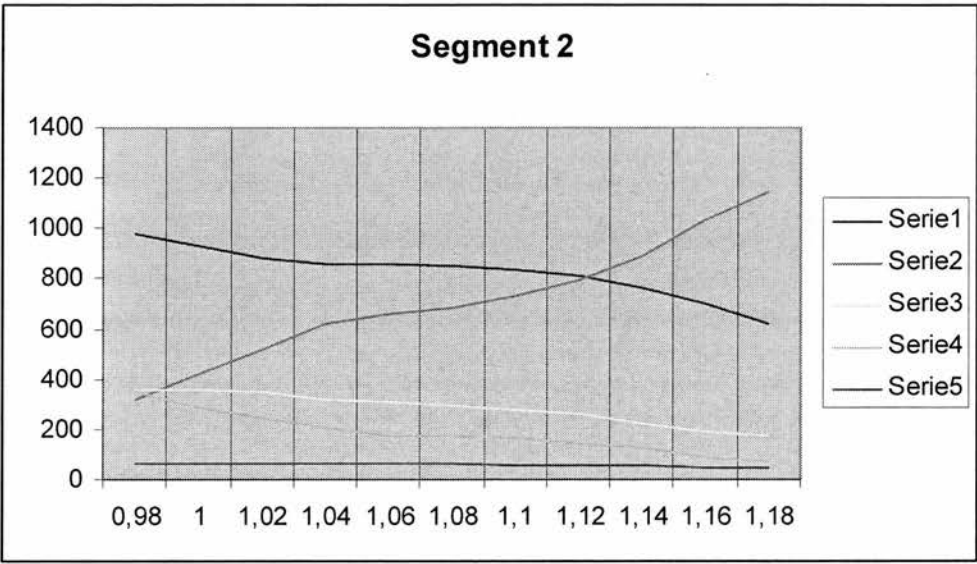


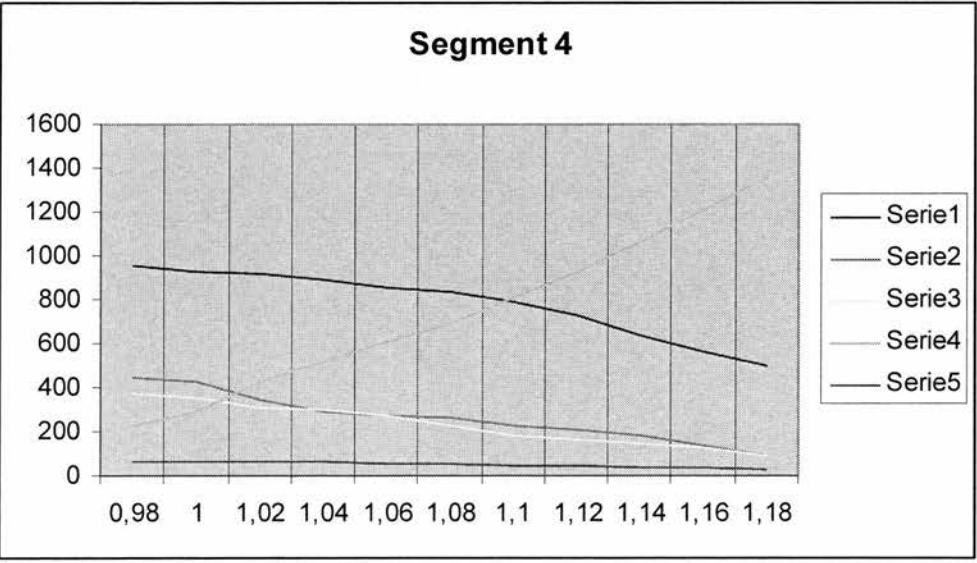
Figure 6.8 illustrates the intervention process results when the corrector parameter is applied to segment 2. It is evident that when forcing segment 2, their customers tends to move to segment 1 and 4.

Figure 6.8 Results from the application of intervention process in segment 2



It will be seen that proportionally increasing the adequacy of individuals vis-à-vis segment 2 quickly empties segment 1 and segment 4. When repeating the procedure with segment 4, the following changes occur (Figure 6.9).

Figure 6.9 Results from the application of intervention process in segment 4



Having applied the intervention process, it is seen that the segments 1, 2 and 4 are merged in one same segment. Then the 5 segments of the classifications have been

reduced to 3 segments to permit comparison with classification Frank 0.454.

Table 6.18 Classification 2 after the intervention process

3. FRANK 0.443			Total segment	% distance buyers/total distance buyers
Segment 1'=1+2+4	Distance buyers	67	1640	85,90%
	Others	1573		
Segment 2'=3	Distance buyers	8	357	10,26%
	Others	349		
Segment 3'=5	Distance buyers	3	66	3,85%
	Others	63		
TOTAL	Distance buyers	78	2063	100%
	Others	1985		

The new classification presents 3 segments. The resulted segment 1' groups the majority of distance buyers which represents the 85.9% of the total. 67 distance buyers is the result of joining the number of distance buyers initially located in segments 1, 2 and 4 before the intervention process.

When comparing the number of distance buyers in each segment from classification 2 (See Table 6.8) with the results from classification 3 (Frank 0.454), we can see that the segment which joins the maximum number of distance buyers is segment 1'. Then, classifications Frank 0.443 is chosen.

6.4.3 Validation

At this stage, the information corresponding to PoO T2 was used. The main goal was to assess LAMDA's predictions. Forecasts made by the unsupervised LAMDA algorithm were then tested with the information provided by V29, which captured whether the customer had bought online at least once since July 2001 until December 2003. As previously mentioned, a first validity was done in January 2002, but the real online customers were just 93. Therefore, the PoO T2 was extended to December 2003. At that time, online customers had just increased by 10 new online customers.

It must be remembered that, the results were compared and analysed from a marketing standpoint. The main goal was to predict the customers who were going to buy online. As Table 6.19 shows, there was the possibility to make two mistakes, but according to our goal the main error was not to identify a real online customer.

Table 6.19 Interpretation of results based on a marketing standpoint

		LAMDA's model forecasts	
Real Behaviour		Online purchaser	Pure Off line Purchaser
	Online Purchaser	SUCCESS	ERROR
	Pure off line Purchaser	ERROR	SUCCESS

As described in the previous experiment, LAMDA has the capability to either assign each customer in just one segment or to assign each customer to each segment according to the GAD to each segment. Based on that, the two possibilities are considered for the validation stage.

Maximum GAD is considered (non overlapping)

The first type of validation did not consider the possibility of overlapping between segments (See Table 6.20).

Table 6.20 Measuring LAMDA's forecasts (no overlapping)

Segment	Number of customers located by LAMDA in each segment	Distance buying clients located in the segment	Number of Real Internet buyers within the segment (from V29)	
1'	1640	67	76	73.78% (76/103)
2'	357	8	19	18.44% (19/103)
3'	66	3	8	7.76% (8/103)
TOTAL	2063	78	103	100%

Segment 1' was proposed by LAMDA to be the one which joined the future online customers. Results show that the majority of the real online customers (73.78%) were located in this segment.

Forecasting accuracy was validated. As shown (See Table 6.21) LAMDA success rate was 73.78%. LAMDA identified 76 potential candidates among the 103 real online customers. However LAMDA also failed to recognize 27 real online customers as potential candidates.

Table 6.21 Interpretation of numerical results (no overlapping is considered)

	LAMDA's online purchasers	LAMDA's pure offline purchasers
Real online buyer	73.78%	26.20%

However, results in Table 6.21 were based on the maximum GAD in one segment but not with the possibility to locate the same customer in more than one segment.

Multiple GADS were considered (overlapping)

The capability of the model to assign each customer to each pattern of behaviour with its own GAD was taken into account for this second validation. In particular, for testing the results based on this overlapping of behaviours, customers with similar adequacy degree in more than one segment were assumed to belong to both segments simultaneously. As mentioned in the previous experiment, a similar adequacy degree is considered when the difference between the GAD of the same customer in each segment is inferior than 0.015. Taking this simultaneity into consideration, a new matrix resulted.

Table 6.22 Measuring LAMDA's forecasts (overlapping)

Segment	Number of customers located by LAMDA in each segment	Distance buying clients located in the segment	Number of Real Internet buyers within the segment (from V29)	
1'	1640	67	90 (76+12+2)	87.37%(90/103)
2'	357	8	19	18.44% (19/103)
3'	66	3	8	7.76% (8/103)
TOTAL	2063	78	103	100%

Table 6.22 shows that 12 of the real internet customers located in segment 2' presented a high GAD to segment 1'. Therefore, it can be interpreted that these 12

customers had the same possibility to behave as online customers as offline customers. They were assumed to possibly behave differently, according to the situation. Just two of the real internet buyers located in segment 3' behaved in the same way that the customers in segment 1'. Based on that, the numerical interpretations of results when overlapping was considered are summarized in Table 6.23.

Table 6.23 Interpretation of numerical results (overlapping is considered)

	LAMDA's online purchasers	LAMDA's pure offline purchasers
Real online buyer	87.37%	26.20%

The number of customers who had bought online at SUPSA was still low (103), but at that time, official Lleida’s Internet figures not only showed a 23% of Internet access between the citizens of Lleida but also a 10% purchasing rate. However, based on the real number, it is demonstrated that LAMDA forecasting accuracy increases when the multiple GADs are considered.

Summary

This experiment is the first research study that has used unsupervised learning techniques to forecast online purchasing based on in-store data. The most crucial stage when implementing the unsupervised forecasting model is the experts’ participation. It is evident that the LAMDA’s unsupervised learning approach is more human-expert dependent than the supervised approach.

The experiment demonstrates that a predictive variable is required. There is no way to track the classification as LAMDA is a black box. Consequently, it is essential to count on a relevant variable that, although it is not exactly the same predictor, it has a high relationship with it.

Furthermore, the possibility to forecast online purchasers also identified whether the online customers are going to continue buying off line (the overlapping cases) is

demonstrated in the experiment as not only the extreme behaviours are identified but also the ambiguous ones. Based on this, the interpretation stage is a key point. Each customer is assigned to one segment. However, there is also the possibility to assign each customer to each segment. This capability of the model has a managerial advantage which consists of interpreting the results by a non-crisp point of view. All the customers who do not always behave in the same way may be identified. Therefore, the theoretical topic that the customer is likely to behave differently according to the situation (See Chapter 2) is currently quantified in this experiment.

Despite the results being quite encouraging for future research, a high number of online customers would be needed to have a realistic measure of the forecasting success ratio of the LAMDA's unsupervised forecasting model.

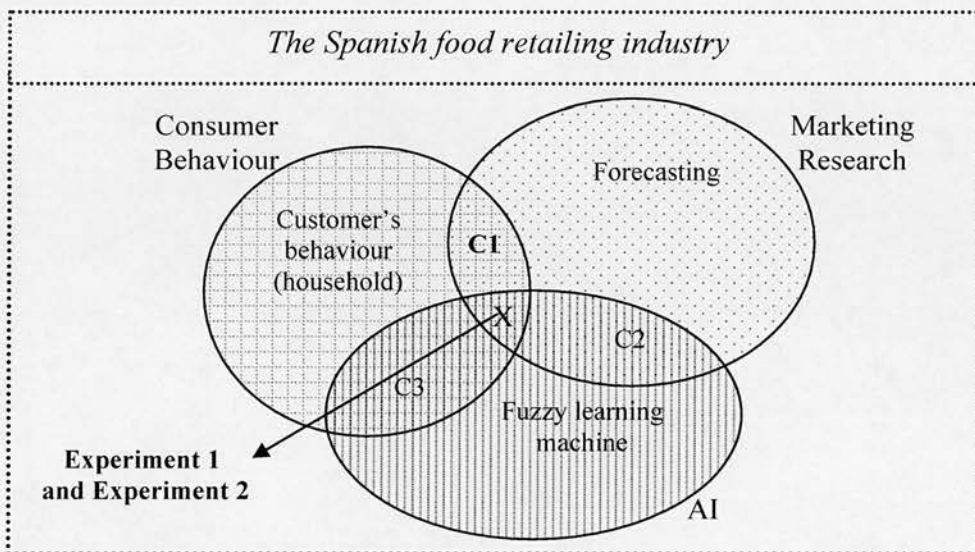
CHAPTER 7

Conclusions

Introduction

This thesis demonstrates the possibility of developing research synergies when connecting 3 different areas of research namely consumer behaviour, marketing research and Artificial Intelligence. Despite the fact that these three areas of research are extensive and when analysing the potential links between them, a wide number of combinations arise, forecasting customers' behaviour using a fuzzy learning technique has been selected to be the triple combination studied in this thesis. The Spanish food retailing industry has been chosen to demarcate the framework of the research.

Figure 7.1 Scope of the thesis



As illustrated in Figure 7.1, two experiments have been carried at the intersection between consumer behaviour, marketing research and AI. Experiment 1 has provided a model to forecast the customers who are likely to defect when the competitor opens a new store. Based on historic shopping data, the fuzzy algorithm applied in the experiment has learnt from past patterns of behaviours to anticipate future behaviours

in similar situations. Although the learning and recognition stages are automatically performed, professional expert opinions have been essential when delimiting the boundaries between the existing patterns of behaviour.

Experiment 2 is the first research study that has used unsupervised learning techniques to forecast online purchasing based on in-store data. The experiment has demonstrated that a predictive variable is required before the implementation of the model. Consequently, it is essential to count on a relevant variable that, although it is not exactly the same predictor, it has a high relationship with it.

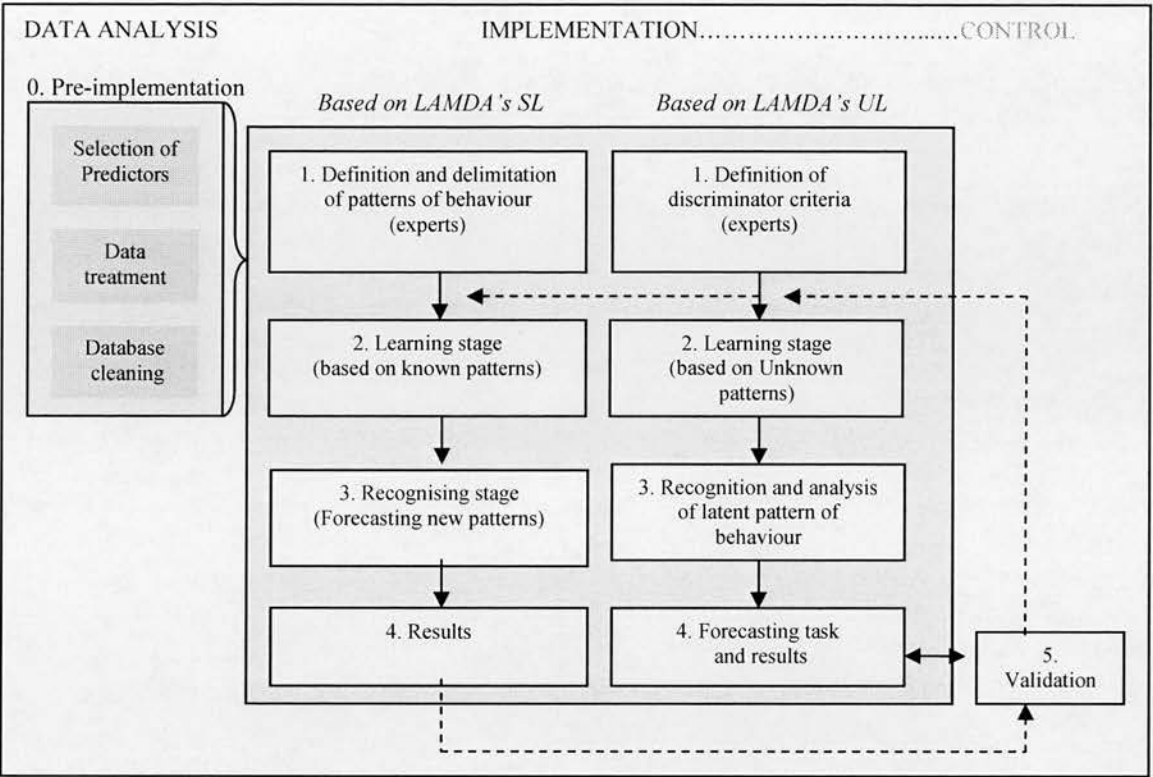
Conclusions from the experiments are drawn. It is also important to note that considerations related to the three intersections areas (C1, C2 and C3) illustrated on the previous figure have been taken into account as well. Although this thesis is multidisciplinary, the main emphasis and focus is in Marketing. Therefore, this chapter is mainly focused on the C1 intersection, as it corresponds to the intersection between consumer behaviour and marketing research. Based on that, and from the food retailing industry standpoint, conclusions are divided into 4 sections and according to Figure 7.1 intersections. They are ordered as follows:

- Empirical research's conclusions (intersection X),
- Consumer behaviour and marketing research (intersection C1),
- Consumer behaviour and Artificial Intelligence (intersection C2),
- Marketing research and Artificial Intelligence (intersection C3).

7.1 Empirical research's conclusions (Intersection X)

Based on the experiments, a dynamic forecasting model based on LAMDA's supervised and unsupervised learning has been built. Figure 7.2 shows the structure of the model.

Figure 7.2 Forecasting model



In respect of data analysis the selection of the predictors was the most critical part. At this stage, expert input was needed to share knowledge and experience related to the specific problem. A knowledge elicitation process was undertaken. For example, in Experiment 1, experts had to list all the variables that they would take into account when determining customer defections. In experiment 2, they had to list the variables which would help them to determine shopping convenience. It is important to note that expert knowledge is based on concepts which are ambiguous and imprecise. Moreover, the same concept may have different meaning depending on the expert's point of view. Therefore, finding out, understanding, labelling, defining and looking for this information in the database was a significant task. In addition, some of the variables which were relevant for the expert were not accessible. Some of them were transformed and substituted, for example the information relating to the means of transport resulted from the address variable. Others simply did not exist, for example the time a customer spent inside the store.

Related to the implementation, the aim of the forecasting model was to imitate the expert's decision process. Despite the fact that experts provided different final decisions, the research was focused on the stages of the decision process and the relevance of the information suggested by them, which were virtually the same. The main difference was related on how they interpreted the information. The main cause which resulted in different final decisions was the way the same information was interpreted by different experts. In order to solve this issue about multiple interpretations of data, they were asked to indicate the information which they thought had to be precise and the data that could be more ambiguous, although relevant. For instance, total purchases were chosen as precise data. However, frequency of visits did not have to be measured as a specific number. An interval was enough.

Despite the fact that the data analysis and control stages are similar from both experiments, it has been demonstrated in Chapters 5 and Chapter 6 that the implementation stage of the supervised or unsupervised learning presents dissimilar peculiarities. Unlike unsupervised learning, supervised learning is based on known patterns of behaviour which make the forecasting task more straightforward. On the other hand, unsupervised learning has shown the necessity to establishing the discriminator criteria (See Figure 7.2) which determines which classifications fulfil the suitable forecasting skills to be selected as the final one. It has been demonstrated that these criterion have a direct affect on the final results.

Related to the control stage, a key conclusion can be drawn which has to do with the interpretation of the results. Both experiments have shown that when forecasting customers' behaviour, each customer is located to the segment which presents the maximum adequacy degree. However, a multiple assignation of the same customer to more than one segment is also possible by dealing with the multiple GADs provided by LAMDA's approach. Each customer is allocated to each existing pattern of behaviour with an adequacy degree. The capability to locate each customer to each existing segment provides a more realistic way to interpret the results. From the validation stage a higher success rate has been achieved when the multiple GADs is

accepted rather than when assigning a customer to just one segment. Customers' behaviour is not always black or white. When considering the possibility that a customer may behave in different ways, the forecast results have proved to be more successful. Furthermore, the possibility to assign an adequacy degree to each segment is demonstrated to provide a better interpretation of the reality as not only the extreme behaviours are identified but also the ambiguous ones. Then, we can conclude with support for theoretical statement that the customer behaves differently according to the situation but also that the possible patterns of behaviour that a customer performs in both situations considered by the experiments may be identified and quantitatively measured. Based on that, in Experiment 1 each customer is assigned to 'Total defectors' and 'Stayers' with an adequacy degree. In Experiment 2, each customer is assigned to 'On-line purchaser' and 'off line customers'. This multiple assignation leads to managerial implications explained at the end of next section (7.2).

7.2 Consumer behaviour and Marketing research (C1)

The intersection in Figure 7.1 of consumer behaviour and marketing research results in several conclusions which can be drawn.

Firstly, the two experiments carried out in this thesis have provided a specific way to forecast customer defections and online purchases. As explained in Chapter 3, the forecasting process is composed of a set of stages. The main forecasting stages (Armstrong, 2001) are listed in Table 7.1. The focus of the research problem and methods selected to carry out the empirical research are illustrated. Then, as summarised in Table 7.1, Chapter 5 and Chapter 6 have demonstrated the capability to forecast customer's (household) future behaviour from the secondary data collected from internal company's databases. Supported by judgmental forecasting and LAMDA's quantitative forecasting method, the implementation has been carried out. Moreover, the forecasting success has been assessed by comparing the forecast with the reality.

Table 7.1 Type of forecasting process carried out

Stages	Selected for the experiments
Formulating the problem	Customer's (household) future behaviour
Obtaining Information	Secondary data collected from internal company's databases (scanners systems and loyalty cards)
Selecting and Implementing Forecasting methods	Judgmental forecasting (experts) and Quantitative forecasting methods (LAMDA)
Evaluation forecasting method	Criteria: Realism

Secondly, an overview of studies of consumer behaviour has been carried out before reconsidering the possibility to adjust some of the traditional consumer research fundamentals to the retailing field, in particular food retailing. When referring to supermarket purchases, the consumer behaviour fundamentals are too general. There is a need to adapt the consumer fundamentals to the food retailing framework as suggested in this thesis. Accordingly, some adjustments have been proposed. Table 7.2 shows these adjustments previously explained in section 2.7 when developing what it is called ‘customer behaviour research’.

Table 7.2 Adjustments of consumer behaviour research and customer behaviour research

	Consumer Behaviour research (cognitive approach)	Customer behaviour research (behavioural approach)
Subject of research	Consumer	Customer (regular shopper) Household
Research goal	Understanding and anticipating consumer behaviour	Understanding and anticipating customer behaviour
Prior Data interest	consumer insights (primary data)	Customers behaviour (secondary data)
Decision making model	Based on product choice	Based on store choice and product choice
Main Key stages	Search/pre-purchase evaluation/ consumption	Purchase and post-purchasing evaluation
Main standpoint	Manufacturers	Retailers

Customer behaviour research fundamentals explained in Chapter 2 have been assumed within the thesis. Despite the fact that traditionally consumer behaviour has been analysed from a cognitive approach, the opportunity to analyse customer

behaviour from a behaviourism viewpoint has been presented. Customer behaviour research is focused on understanding and anticipating the way and reasons why people shop, purchase or buy. Therefore, from a retailers' perspective, the main research focus of analysis is the regular shopper instead of the consumer. As far as the decision making model is concerned, repeated purchases in groceries does not imply a complex decision process but a decision making model resulting from a trial and error learning process. Based on that, the prior data required to understand this type of shopping behaviour is considered to be secondary data, captured from past behavioural data. Behavioural data captures the information of the customer related to the purchase and post purchase stages. Within groceries, the seven stages of the decision process described by Engel, Blackwell and Miniard (1995) are virtually reduced to two, which include purchase stage and post purchase stage. The analysis of the purchase act (shopping behaviour) and customer tracking (post purchase evaluations) are critical tasks for the retailers when attempting to enhance the customer-supermarket relationship.

Thirdly, and related to the previous one, a reconsideration of the types of decision making model found in the literature is proposed. Consumers and customers take decisions. However, the decision process that a customer carries out when buying groceries is completely different from the decision process when buying for instance a car. Moreover, the type of decision making model when selecting a store does not have to be the same as the type of decision making model when selecting a product in that store, or the other way round. Different types of consumer decision making are found through the literature. Depending on the time and energies spent when purchasing a product or selecting a store, the type of decision making may vary. However, both the store choice and product choice are analysed separately. It is firmly believed that a final purchase is a result of the combination between a type of decision making when choosing the store, and a type of decision making when choosing the product (s), and vice versa as well. Then, the necessity to find a range of integrated types of decision making has been suggested and a double choice matrix (See Figure 7.3) have been provided describing the existing duality of decision making models, particularly in groceries. As illustrated in the diagonal of

Figure 7.3, the types of customer decision making models coincide when selecting the product or the store.

Figure 7.3 Typology of customer decision making in groceries

<div>Product Choice</div> <div>Store Choice</div>	Complex Decision (extent decision making and high involvement)	Brand Loyalty (habit and high involvement)	Variety Seeking (extent decision making and low involvement)	Inertia (habit and low involvement)
Complex Decision (extent decision making and high involvement)	Complex Decision			
Brand Loyalty (habit and high involvement)		Brand Loyalty		
Variety Seeking (extent decision making and low involvement)			Variety Seeking	
Inertia (habit and low involvement)				Inertia

However, it has been explained that groceries purchases are commonly characterised to be repetitive or eventually purchases based on habit, as in shopping trips and in the supermarket basket purchased. In addition, implicit in repetitive purchasing is the assumption that customers learn from past experience and buy the product or/and at the store what most matches expectations. When this happens, information search and alternatives evaluation are limited or non existent since the customer has decided to buy almost the same basket again. Therefore, the decision making is considerably simplified. Then, although complex decision may appear (for instance when selecting a supermarket brand for the first time) and also variety seeking may become visible, (when there is a wide evaluation of store brand due to a unsatisfactory experience, for example), two most common types of customer decision corresponding to store choice are suggested and assumed, which are brand loyalty and inertia (See Figure 7.3). Furthermore, there are some products which reappear every time the supermarket basket is filled. Some purchases are made again and again. If we look at the items purchased on the supermarket we see that we have bought the same product and the same brand many times before. Based on that, brand loyalty is also possible in product choice that frequent items require ‘little conscious attention’ to be purchased. As illustrated in Figure 7.3, shadowed cells

show the typology of customer decision making when groceries are purchased by a regular shopper. Based on that, 6 out the initial 16 scenarios resulted from merging store choice and product choice decisions are recommended as the types of customer decision making model when referring to supermarkets.

Fourthly, the chameleonic consumer (customer) concept is not new (e.g. Dubois, 1994). The theory that customers behaves differently according to the situation is not recent. However, from a quantitative analysis research point of view, the fact that customer behaves differently according to the situation has been difficult to measure. As mentioned, marketing research aims to turn information into knowledge to support managerial decisions. However, most of the marketing research software tools are based on Aristotle's logic. According to Zadeh (1973), digital computers failed when trying to deal with the reality of human thinking and behaviour. Consequently, he proposed an approach which shifted from the precision and statistical formalisms towards the partial truth. Dealing with the partial truth has been demonstrated to be more realistic when forecasting. A representation and measurement of the possible types of shopping behaviours performed by the same customer is demonstrated in this thesis. From this thesis point of view, each customer belongs to each pattern of behaviour provided by the model in a certain adequacy degree (GAD). Then, GADs are the key point when measuring the theoretical concept of the chameleonic customer. Sections 5.4.3 and 6.4.3 have demonstrated that when considering customers' GADs, the forecasting task is more successful. Accordingly, targeting decisions became more realistic.

From a managerial perspective, the two experiments have introduced a new way to interpret the results to support decision making in marketing. Particularly, the fact that each customer presents a specific adequacy degree to each segment affects the traditional way of targeting. Graphically, Figure 7.4 and Figure 7.5 illustrate an example for this new managerial approach.

Figure 7.4 Example: Adequacy degree of each customer to the pattern of behaviour 'Stayer'

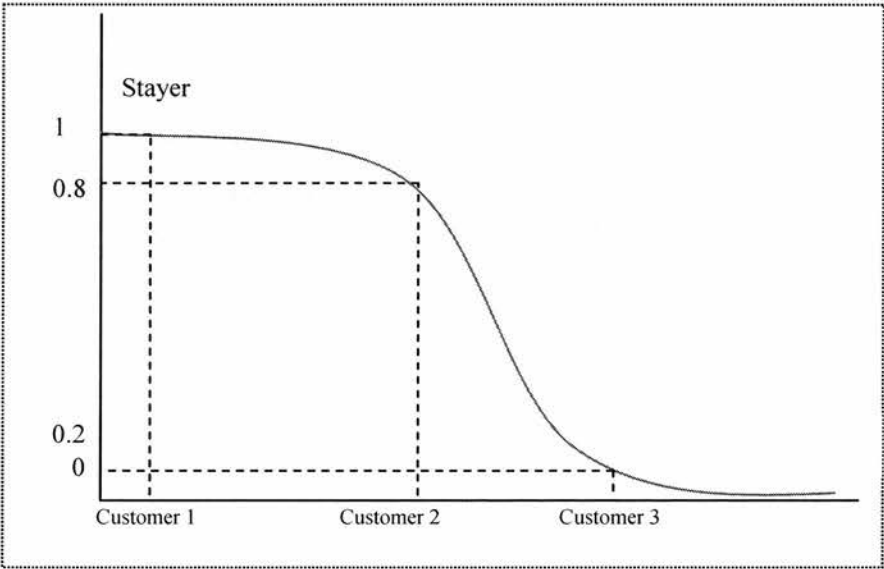
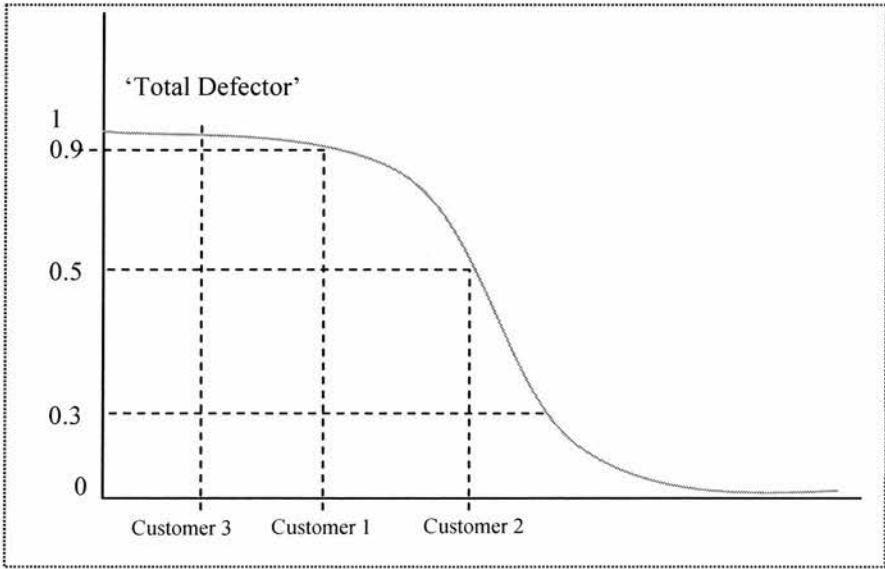


Figure 7.5 Example: Adequacy degree of each customer to the pattern of behaviour 'Total Defector'



As seen in Figures 7.4 and 7.5 each customer has a specific adequacy degree to each pattern of behaviour ('Stayers' and 'Total defectors'). When analysing customer 3, we can see that he/she has a low adequacy degree in 'Stayers' and a high adequacy degree in 'Total defectors'. Then, it is clear that he is likely to behave as a 'Total

defectors'. On the other hand, there is Customer 1 who presents a high adequacy degree in both patterns of behaviour. Although his/her maximum GAD is in 'Stayers' (See Figure 7.4), he/she also has a considerable GAD to defectors (See Figure 7.5). Finally, customer 2 has been assigned to 'Stayers' with a 0.8 GAD, and to 'Total Defectors' with a 0.5 GAD. The possibility to be able to find out about the ambiguous customers has many positive implications for managers. Mainly, marketing managers are able to be more realistic when targeting marketing actions in order to retain their customers. Then, the marketing action is not going to be focused to the most valuable customers who have been assigned in the 'Total defector' segments, but also to other customers that, even having a high adequacy degree in 'Stayers' also presents a considerable GAD in 'Total defectors'. They are ambiguous customers, who may behave differently according to their personal situations and to whom the company is interested in retain as well. Based on the examples in the two previous figures, all the customers should be targeted. However, this is not the point. The most important fact of this new approach is that according to the budget and resources, the manager may select the GAD which will delimit the marketing action. For instance, when targeting all the customers who show a GAD of at least 0.7 at the 'Total defectors', he is selecting customer 3 and customer 1, but not customer 2. Based on that, customer 2 is not considered a risky loss, and he is not going to receive a specific direct marketing action aimed to stay with the company.

7.3 Consumer behaviour and Artificial Intelligence (C2)

The existing complementarities between consumer behaviour and Artificial Intelligence have been demonstrated in this thesis. On one hand, it is a shared accepted statement that understanding and forecasting consumer behaviour is a difficult task. For instance, when considering the most complex model of consumer decision model, several factors appear such as attitudes, intentions, values, opinions, memory, and experience. Time and situation also may influence the way consumers (customers) consume or purchase. These factors are ambiguous, uncertain and imprecise. Then, when analysing or predicting customers' behaviour automatically, an interval of flexibility is needed to cope with the ambivalence of their behaviour.

Consumers are complex. It has been demonstrated that fuzzy logic provides approximate but consistent solutions to complex problems, where numerical data usually are noisy and incomplete and the linguistic information is imprecise and vague (Tchamova and Semerdjiev, 2002).

Furthermore, it has been highlighted that customers' patterns of behaviour are the focus of learning. Customers are not simple robots. People tend to behave differently according the situation so their patterns of behaviour are not always certain and rigid. To deal with this dynamic, uncertain and complex aspect, fuzzy logic is considered as it is based on the concept of partial truth.

Then, we really believe and it is proposed in this thesis to apply fuzzy logic theory when understanding customers. Moreover, technological advances allow companies to collect vast quantities of information about the customers. The information is often ambiguous and incomplete. AI is able to deal with this type of information.

7.4 Marketing research and Artificial Intelligence (C3)

These two experiments represent two more cases to add to the marketing research trend which is focused on applying advanced new tools to solve traditional problems. Mention that LAMDA's software was initially developed to work in industrial processes control and to environmental systems simulation. In fact, some adjustments have been required for its proper application to the marketing field. In particular, a set of criterion has been established and specific solutions set up for each of the experiments. These adjustments have been described in Chapters 5 and 6.

A review of other applications has been provided in Chapter 3. As explained in section 3.3.5, there is an ongoing discussion whether statistical techniques are better than artificial intelligence techniques when forecasting. It has been reviewed that some researchers enhance the advantages of the AI tools, others support the traditional techniques and many highlight the complementarities between artificial intelligence and statistics. In fact comparing LAMDA with another statistical technique was not the aim of this thesis. It has been tested that LAMDA can deal

with qualitative and quantitative data simultaneously. It has been experienced that LAMDA may be used as forecasting tool. However, concluding whether LAMDA performs better than some statistical model is not possible and it was not our research goal. As mentioned in section 4.6, LAMDA is similar to the basic model of neural networks (NN). When considering the most quoted advantages and disadvantages of NN provided by Vellido, Lisboa and Vaughan, (1999) and applying them to LAMDA's performance, some conclusions can be drawn. In general terms, we can conclude that LAMDA, as a fuzzy learning technique similar to NN, has shown the majority of advantages and disadvantages frequently analysed when studying neural nets.

7.5 Limitations and Further Research

Forecasting customer's shopping behaviours applying a fuzzy learning technique has its advantages and its limitations. As far as the *limitations* are concerned, despite the fact that behavioural data may provide important information about the current customers, it is important to note that it is incomplete. The collected data from the loyalty cards and scanners systems provides information about the customer and the store, but not what the shopping behaviour is when the customer leaves the store. Then, it is clear that these scanner systems and loyalty card schemes shows a partial description of the real word as they do have not much to say about the total market and competitor activity (Uncles, Ehrenberg and Hammond, 1995). For instance, in the food retailing sector, many consumers are variety seekers (McGoldrick, 2002) who combine their supermarket basket across more than one store, and obviously, each supermarket just captures part of the whole customer behaviour. A part from this partial representation of the reality, there are other customers who regularly come to the stores but they are not interested in the loyalty card, or in the way the supermarket tracks their behaviour. This type of customers also demarks the limits of the research.

In reference to Experiment 1, a considerable limitation was found. As explained in Chapter 5, LAMDA's supervised learning approach learnt from customer's reactions

when a competitor opened a store in Trempe to forecast customer's future behaviour when the same competitor opened another store in Flix. However, the learning stage may be performed properly when the scenarios (Trempe and Flix) are similar. After analysing the data from each village we were able to realise that Trempe and Flix were not as similar as the experts initially considered. This dissimilarity between scenarios affected the final results. Based on that, we conclude that statistical procedures may be helpful to test the similarity between the contexts before starting the implementation stage of the model. Particularly to Experiment 2, the main limitation was related to the validation corpus. As mentioned in Chapter 6, the number of customers who had bought online was still insignificant to take strong decisions, specifically to design suitable marketing actions.

The interpretation of the available data, in particular coming from the loyalty cards is not easy (Uncles, Ehrenberg and Hammond, 1995; Baron and Lock, 1995), not only for the vast amount of data collected but also because expert opinion is needed. As shown in sections 5.4 and 6.4, expert knowledge directly affected LAMDA's model of forecasting, as the learning stage imitates the way experts learn. Furthermore, some of the information considered relevant by the expert was not available, because of its cost or because at that time it did not exist. For instance, time of shopping would be one example of existing data but inaccessible. Depending on the moment of purchase, the shop might be overfull, the personnel may be tired, the aisle may be relatively empty, etc. Therefore, this variable may lead to service quality differences across time of day (Buckinx and Van den Poel, 2005).

Additionally, input data used in the experiments are related to shopping behaviours but not customer's attitudes, beliefs or emotions. There was no information in the company's internal databases about customers' personal opinions of the supermarket. For instance, and affecting experiment 1, knowing customers who had complained at least once, would have also helped to forecast customer attrition. According to Roos (1999), merely complaining does not cause switching but the response to it has a decisive effect on customer reactions.

In reference to *further research*, there are two possible future lines of research that could be developed. The first is directly related to the two experiments. The other future line of research is more general and aims at continuing to demonstrate the synergies between the three main topics introduced in Figure 7.1.

Related to the further research on the experiments, some aspects are proposed. Firstly, it would be interesting to test the same problem of research but adding other information such as primary data (e.g. personality traits, customers' preferences, customers' opinions and values) or external secondary data (e.g. TNS panels). In addition, data captured from new mechanical devices including galvanometers, tachistoscope, eye cameras and audiometer would also be interesting. Then, analysing the possibility to forecast customers' behaviour combining both types of data could be undertaken. Each type of collecting method provides a specific type of information. As has been explained, LAMDA can deal with qualitative and quantitative data. Based on that, combining these data with the input data applied in Experiments 1 and 2 would help to finding possible improvements to forecast customer defections and online purchases.

Secondly, LAMDA is able to learn from past patterns of behaviour to forecast new patterns in similar context. The possibility to use Trepmp and Flix customers and learn from them to forecast what customers from other villages would do in a similar situation would also help to check the retro-feeding capability of the forecasting model.

Thirdly, applying the same experiments to different industries rather than Spanish food retailing industry would be interesting. There are an increasing number of companies which invest in similar schemes to loyalty cards or scanners systems to collect data from their customers. Retailers, airlines, hotels, banks, car companies, telecommunication companies, media planning companies, fashion stores, universities and ecommerce platforms all collect this type of data. The possibility to apply the same model to other businesses is suggested as a future research line.

Finally, comparing LAMDA's forecasting model to other existing techniques would be a possible future direction of research. Chapter 3 has shown the difficulty in measuring the accuracy of a forecasting technique. Although one possible way to check the forecasting success is comparing the forecasts with the reality, it is also interesting to compare LAMDA's results to other technique, for example to a neural network, cluster analysis or logistic regression.

Related to more general future research lines, noting again that this thesis has shown the possibility of developing synergies when connecting 3 different areas of research namely consumer behaviour, marketing research and AI. We believe that there is a wide set of triple combinations relating to these fields. In addition, many experiments could be developed in different industries such as banks, fashion retailers, airline companies, online stores and music stores. Some examples are proposed:

- Segmenting customer shopping frequency using a fuzzy neural networks in Spanish music industry,
- Forecasting organic food consumption preferences using LAMDA in the English food retailing industry,
- Segmenting brand choice using LAMDA in the Italian fashion retailing industry,
- Segmenting teenager's behaviour using genetic algorithms in online music industry,
- Identifying customer's shopping behaviour using web tags and multiagent systems in online book shop stores,
- Forecasting consumer brand choice using LAMDA in the Spanish food retailing industry,
- Forecasting price elasticity using fuzzy neural networks the Spanish airline companies, etc.

Moreover, not only associations between these areas of research are possible, but also synergies between other marketing management areas which deal with ambiguous, uncertain and incomplete information. Based on that, a list of possible future experiments is also suggested:

- Measuring the influence of brand awareness in consumer behaviour using LAMDA,
- Forecasting media planning success ratio using Neural Networks,
- Measuring the influence of publicity in a product category using LAMDA,
- Recognising new possibilities of product design using LAMDA,
- Forecasting product sales using fuzzy neural networks,
- Measuring Brand store awareness using LAMDA,
- Segmenting store behaviour patterns using fuzzy neural networks, etc.

To conclude, it is important to note the main similarity between the experiments developed in this thesis and the previous suggested examples. All of them share the same condition which is the existing patterns of behaviour. It does not matter whether the patterns of behaviour are based on customers, consumers, stores, products or brands as long they are well defined with relevant information. It does not matter whether the type of variables which compose the patterns are quantitative or qualitative, primary, secondary or both, as long as they are shared by experts in the field. It does not matter whether the tasks carried out when analysing the data are forecasting, segmenting or identifying , as long as the information and the object of research is clear and there are experts open to share their knowledge and interest in working in a multidisciplinary environment.

BIBLIOGRAPHY

- AAAI. "Cognitive Science." Web page, [accessed December 2004]. Available at <http://aaai.org/AITopics/html/cogsci.html>.
- Aaker, D. A. and Jacobson R. "The Sophistication of Naive Modeling." *International Journal of Forecasting* 3 (1987): 449-451.
- Aamodt, A. and Plaza E. "Case-Based Reasoning: Foundational Issues, Methodological Variations and Systems Approaches." *Artificial Intelligence Communications* 7, no. 1 (1994): 39-59.
- Achabal, D. D. and McIntyre S. H. "Information Technology Is Reshaping Retailing." *Journal of Retailing* 63 (1987): 321-325.
- Adams, F. G. *The Business Forecasting Revolution*. New York: Oxford University Press, 1986.
- AECE. "Resumen del Estudio sobre comercio electrónico B2C." Web page, 2003 [accessed December 2004]. Available at <http://www.aece.org>.
- Agrawal, D. and Schorling C. "Market Share Forecasting: An Empirical Comparison of Artificial Networks and Multinomial Logit Model." *Journal of Retailing* 72, no. 4 (1996): 383-407.
- Aguado, J. C. "A Mixed Qualitative-Quantitative Self - Learning Classification Technique Applied to Simulation Assessment in Industrial Process Control." Universitat Politècnica de Catalunya, 1998.
- Aguilar-Martin, J. and López de Mántaras R. "The Process of Classification and Learning: the Meaning of Linguistic Descriptors of Concepts." *Approximate Reasoning in Decision Analysis* (1982): 165-175.
- Aguilar-Martin, J. and Piera N. "Les Connectifs Mixtes: de Nouveaux Opérateurs d'Association des Variables dans la Classification Automatique avec Apprentissage." In *Data Analysis and Informatics*, E. Diday et al., Elsevier Science Publishers, 1986, 253-265.
- Ahmad, R. and Buttle F. "Customer Retention Management: a Reflection of Theory and Practice." *Marketing Intelligence & Planning* 20, no. (3) (2002): 149-161.
- Aiken, M. "Artificial Neural Systems as a Research Paradigm for the Study of Group Decision Support System." *Journal of Retailing* 6, no. 4 (1997): 373-382.

AIMC. "Audiencia en Internet (abril/mayo 2001)." Web page, November 2001 [accessed December 2003]. Available at <http://www.aimc.es>.

Ainscough, T. L. and Aronson J. E. "An Empirical Investigation and Comparison of Neural Networks and Regression for Scanner Data Analysis." *Journal of Retailing and Consumer Services* 6, no. 4 (1999): 205-217.

Ajzen, I. and Fishbein M. "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research." *Psychological Bulletin* 84 (1977): 888-918.

———. *Understanding Attitudes and Predicting Social Behavior*. New Jersey: Prentice Hall, 1980.

Ali, J. and Rao C. P. "Micro-Market Segmentation Using a Neural Network Model Approach ." *Journal of International Consumer Marketing* 13, no. 2 (2001): 7-27.

Allen, J. J. Iacono W. G. and Danielson K. D. "The Identification of Concealed Memories Using the Event-Related Potential and Implicit Behavioral Measures: A Methodology for Prediction in the Face of Individual Differences." *Psychophysiology* 29 (1992): 504-522.

Alon, I. Qi M. and Sadowski R. J. "Forecasting Aggregate Retail Sales: A Comparison of Artificial Neural Networks and Traditional Methods." *Journal of Retailing and Consumer Services* 8, no. 3 (2001): 147-156.

Alpaydin, E. *Introduction to Machine Learning*. Massachusetts: Massachusetts Institute of Technology, 2004.

AMA. "Marketing Definitions: A glossary of marketing terms." Web page, 2004 [accessed November 2004]. Available at <http://www.ama.org>.

Anderson, E. W. and Sullivan M. W. "The Antecedents and Consequences of Customer Satisfaction for Firms." *Marketing Science* 12, no. 2 (1993): 125-143.

Anderson, E. W. and Weitz B. "Determinants of Continuity in Conventional Channel Dyads." *Marketing Science* 8, no. Fall (1989): 310-323.

Anderson, J. A. *An Introduction to Neural Networks*. Massachusetts: The MIT Press, 1995.

Anderson, J. R. *The Architecture of Cognition*. Massachusetts: Harvard University Press, 1983.

Anil, K. J. Duin R. P. W. and Mao J. "Statistical Pattern Recognition: A Review." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, no. 1 (2000): 4-36.

Ansuj, A. P. Camargo M. E. Radharamanan R. and Petry D. G. "Sales Forecasting Using Time Series and Neural Networks." *Computers and Industrial Engineering* 31, no. 1/2 (1996): 421-424.

Anthony, M. and Bartlett P. L. *Neural Network Learning: Theoretical Foundations*. United Kingdom: Cambridge University Press, 1999.

Araya, S. Silva M. and Weber R. "A methodology for web usage mining and its application to target group identification." *Fuzzy Sets and Systems* 148, no. 139-152 (2004).

Armstrong, J. S. *Long Range Forecasting: From Crystal Ball to Computer*. New York: John Wiley & Sons, 1985.

Armstrong, J. S. and Brodie R. J. *Forecasting for Marketing*. United Kingdom: International Thompson Business Press, 1999.

Armstrong, J. S. and Collopy F. *Integration of Statistical Methods and Judgment for Time Series Forecasting: Principles From Empirical Research*. New York: John Wiley & Sons, 1998.

Armstrong, J. S. ed. *Principles of Forecasting: Handbook for Researchers and Practitioners*. Norwell, MA: Kluwer Academic Publishers, 2001.

Armstrong, J. S. Morwitz V. G. and Kumar V. "Sales Forecasts for Existing Consumer Products and Services: Do Purchase Intentions Contribute to Accuracy." *International Journal of Forecasting* 16 (2000): 383-397.

Ashton, A. H. "Combining the Judgments of Experts: How Many and Which Ones?" *Organizational Behavior and Human Decision Processes* 38 (1986): 405-414.

Assael, H. *Consumer Behaviour and Marketing Action*. California: Kent Publishing Company, 1987.

Athanassopoulos, A. D. "Customer Satisfaction Cues to Support Market Segmentation and Explain Switching Behavior." *Journal of Business Research* 47, no. 3 (2000): 191-207.

Aylott, R. and Mitchell V-W. "An Exploratory Study of Grocery Shopping Stressors." *International Journal of Retail & Distribution Management* 26, no. 9 (1998): 362-373.

Back, B. Irjala M. Sere K. and Vanharanta H. "Competitive Financial Benchmarking Using Self-Organizing Maps." *Paperi Ja Puu -Paper and Timber* 79, no. 1 (1997): 42-49.

- Baesens, B. Viaene S. Van den Poel D. Vanthienen J. and Dedene G. "Bayesian Neural Network Learning for Repeat Purchase Modelling in Direct Marketing." *European Journal of Operational Research* 138 (2002): 191-211.
- Baets, W. R. J. Brunenberg L. and van Wezel M. "Using Neural Network Based Tools for Building Learning Organisations." *Accounting, Management & Information Technology* 8 (1998): 211-226.
- Bagozzi, R. P. "Attitudes, Intentions and Behavior: A Test of Some Hypotheses." *Journal of Personality and Social Psychology* 41, no. October (1981): 607-627.
- Bails, D. G. and Peppers L. C. *Business Fluctuations: Forecasting Techniques and Applications*. New Jersey: Prentice Hall International, 1993.
- Bakos, J. Y. "Reducing Buyer Search Costs: Implications for Electronic Marketplaces." *Management Science* 43, no. 12 (1997): 1676-1692.
- Balakrishnan, P. V. S. Cooper M. C. Jacob V. S. and Lewis P. A. "Comparative Performance of the FSCL Neural Net and K-Means Algorithm for Market Segmentation." *European Journal of Operational Research* 93 (1996): 346-357.
- Baldinger, A. L. and Robinson J. "Brand Loyalty: The Link between Attitude and Behavior." *Journal of Advertising Research* 36, no. 6 (1996): 22-35.
- Bareham J. R. *Consumer Behaviour in the Food Industry: a European Perspective*. United Kingdom: Butterworth-Heinemann, 1995.
- Barlow, H. B. "Unsupervised Learning." *Neural Computation* 1 (1989): 295-311.
- Baron, S. and Lock A. "The Challenges of Scanner Data." *Journal of the Operational Research Society* 46 (1995): 50-61.
- Batchelor, R. and Dua P. "Forecaster Diversity and the Benefits of Combining Forecasts." *Management Science* 41 (1995): 68-75.
- Bawa, K. and Ghosh A. "A Model of Household Grocery Shopping Behavior." *Marketing Letters* 10, no. 2 (1999): 149-160.
- Becker, S. "Unsupervised Learning Procedures for Neural Networks." *International Journal of Neural Systems* 2, no. 1/2 (1991): 17-33.
- Belk, R. W. "An Exploratory Assessment of Situational Effects in Buyer Behaviour." *Journal of Marketing Research* 11, no. May (1974): 156-163.
- . "Issues in the Intention-Behavior Discrepancy." In *Research in Consumer Behavior Vol 1*, J.N. Sheth, ed., JAI press, Greenwich, 1985, 1-34.

- . "Situational Variables and Consumer Behaviour." *Journal of Consumer Research* 2, no. 3 (1975): 157-164.
- Bellenger, D. N. Robertson D. H. and Greenberg B. A. "Shopping Centre Patronage Motives." *Journal of Retailing* 53, no. Summer (1997): 29-38.
- Belt, D. "Neural Networks: Practical Retail Applications." *Discount Merchandiser* October (1993): 9-11.
- Bemmaor, A. C. "Predicting Behavior from Intention-to-Buy Measures: The Parametric Case." *Journal of Marketing Research* 32, no. May (1995): 176-191.
- Bentler, P. M. and Speckart G. "Models of Attitude-Behavior Relationships." *Psychological Review* 86 (1979): 452-464.
- Berrell, J. "Just-in-Time Shoppers." *Retail World* 48, no. 12 (1995): 26.
- Berry, M. J. A and Linoff G. S. *Data Mining Techniques: For Marketing, Sales and Customer Support*. New York: John Wiley & Sons, 1997.
- Berstein, D. A. Penner L. A. Clarke-Stewart A. and Roy E. J. *Psychology*. Boston: Houghton Mifflin Company, 2003.
- Bettman, J. R. *An Information Processing Theory of Consumer Choice*. Massachusetts: Addison-Wesley, 1979.
- Bhattacharya, C. B. "When Customers Are Members: Customer Retention in Paid Membership Contexts." *Journal of the Academy of Marketing Science* 26, no. 1 (1998): 31-45.
- Bigus, J. P. *Data mining with Neural Networks: Solving Business Problems-From Application Development to Decision Support*. Indiana: McGraw-Hill, 1996.
- Blattberg, R. C. and Hoch S. J. "Database Models and Managerial Intuition: 50% Model + 50% Manager." *Management Science* 36, no. August (1990): 887-889.
- Blaylock, J. R. "An Economic Model of Grocery Shopping Frequency." *Applied Economics* 21 (1989): 843-852.
- Bloemer, J. M. M. Brijs T. Swinnen G. and Vanhoof K. "Identifying Latently Dissatisfied Customers and Measures for Dissatisfaction Management." *International Journal of Bank Marketing* 20, no. 1 (2002): 27-37.
- Bloemer, J. M. M. Brijs T. Vanhoof K. and Swinnen G. "Comparing Complete and Partial Classification for Identifying Customers at Risk." *International Journal of Research in Marketing* 20 (2003): 117-131.

Blundin, D. "How to Extract Value from Your Collected Data." *Automatic I.D. News* 14, no. 8 (1998): 42.

Boden, M. A. *Artificial Intelligence and Natural Man*. New York: Basic Books, 1977.

———. *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press, 1990.

Bolt, G. *Market and Sales Forecasting: A Total Approach*. London: Kogan Page Limited, 1994.

Bolton, R. N. "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction." *Marketing Science* 17, no. 1 (1998): 45-65.

Boone, B. S. and Roehm M. "Retail Segmentation Using Artificial Neural Networks." *International Journal of Research in Marketing* 19, no. 3 (2002): 287-301.

Borowsky, M. "Scoring Puts Up Higher Numbers." *US Banker* 105, no. 1 (1995): 63.

Bowlby, R. *Carried Away. The Invention of Modern Shopping*. United Kingdom: Faber and Faber, 2000.

Breitenbach, C. and Van Doren D. "Value-Added Marketing in the Digital Domain: Enhancing the Utility of the Internet." *Journal of Consumer Marketing* 15, no. 6 (1998): 558-575.

Brodie, R. J. and de Kluyver C. A. "A Comparison of the Short Term Forecasting Accuracy of Econometric and Naive Extrapolation Models of Market Share." *International Journal of Forecasting* 3 (1987): 423-437.

Broersma, M. "AI gets down to business." Web page, January 2001 [accessed January 2004]. Available at <http://news.zdnet.co.uk/internet/webservices/0,39020378,2083916,00.htm>.

Brooks, R. A. "Intelligence without Representation." *Artificial Intelligence* 47, no. 1-3 (1991): 139-159.

Buckinx, W. and Van den Poel D. "Customer Base Analysis: Partial Defection of Behaviourally Loyal Clients in a Non-Contractual FMCG Retail Setting." *European Journal of Operational Research* 164, no. July (2005): 252-268.

Buckinx, W. Moons E. Van den Poel D. and Wets G. "Customer-Adapted Coupon Targeting Using Feature Selection." *Expert Systems with Applications* 26, no. 4 (2004): 509-518.

- Bucklin, L. and Gautschi D. *The Importance of Travel Mode Factors in the Patronage of Retail Centres*. New York: North Holland, 1983.
- Bucklin, R. E. and Sismeiro C. "A Model of Web Site Browsing Behaviour Estimated in Click stream Data." *Journal of Marketing Research* 40, no. 3 (2003): 249-267.
- Bunn, D. and Wright G. " Interaction of Judgmental and Statistical Forecasting Methods: Issues and Analysis." *Management Science* 37 (1991): 501-518.
- Burke, L. I. and Ignizio J. P. "Neural Networks and Operations Research: an Overview." *Source Computers and Operations Research Archive* 19, no. 3/4 (1992): 179-189.
- Burke, R. R. "Technology and the Customer Interface: What Consumers Want in the Physical and Virtual Store." *Journal of the Academy of Marketing Science* 30, no. 4 (2002): 411-432.
- Calder, B. J. *Programme Evaluation and Quality. A Comprehensive Guide to Setting up an Evaluation System*. London: Kogan Page, 1994.
- Capraro, A. J. Broniarczyk S. and Srivastava R. K. "Factors Influencing the Likelihood of Customer Defection." *Journal of the Academy of Marketing Science* 31, no. 2 (2003): 164-175.
- Carpenter, G. A. and Grossberg S. "The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network." *IEEE Computer* 21, no. 3 (1988): 77-88.
- Carson, D. Gilmore A. Perry C. and Gronhaug K. *Qualitative Marketing Research*. London, Great Britain: Sage, 2001.
- Charniak, E. and McDermott D. *Introduction to Artificial Intelligence*. Massachusetts: Addison-Wesley, 1985.
- Chatfield, C. "Neural Networks: Forecasting Breakthrough or Passing Fad?" *International Journal of Forecasting* 9 (1993): 1-3.
- Chiang, W-y. K. Zhang D. and Zhou L. "Predicting and Explaining Patronage Behavior Toward Web and Traditional Stores Using Neural Networks: a Comparative Analysis With Logistic Regression." *Decision Support Systems* (2004): forthcoming.
- Chiger, S. "Consumer Shopping Survey: Part 2." *Catalogue Age* 18, no. 11 (2001): 47-51.
- Church, K. B. and Curram S. P. " Forecasting Consumers' Expenditure: a Comparison between Econometric and Neural Network Models." *International Journal of Forecasting* 12 (1996): 255-267.

- Churchland, P. M. *Matter and Consciousness*. London: MIT Press, 1992.
- Cialdini, R. B. *Influence: Science and Practice*. Boston: Allyn & Bacon, 2001.
- Clemen, R. T. "Combining Forecasts: A Review and Annotated Bibliography, (With Discussion)." *International Journal of Forecasting* 5 (1989): 559-588.
- Cohen, J. B. "Attitude, Affect and Consumer Behavior." In *Affect and Social Behaviour*, B.S. Moore and A.M. Isen, eds., Cambridge University Press, NY, 1990, 152-206.
- Collins, D. C. and Lindley T. "Store Brands and Retail Differentiation: The Influence of Store Image and Store Brand Attitude on Store Own Brand Perceptions." *Journal of Retailing and Consumer Services* 10, no. 6 (2003): 345-352.
- Copeland, J. *Artificial Intelligence: A Philosophical Introduction*. Oxford: Blackwell, 1993.
- Cortjens, M. and Lal R. "Building Store Loyalty through Store Brands." *Journal of Marketing Research* 37, no. (3) (2000): 281-291.
- Cox, E. *The Fuzzy Systems Handbook: a Practitioner's Guide to Building, Using and Maintaining Fuzzy Systems*. Chestnut Hill, Massachusetts: AP professional, 1994.
- Coyne, K. "Beyond Service Fads- Meaningful Strategies for the Real World." *Sloan Management Review* 30, no. Summer (1989): 69-76.
- Crawford, F. "Consumer Relevancy: Connecting With the 21st Century Market." *Global Online Retailing Report* (2000): 40-45.
- Cruz, C. A. *Understanding Neural Networks: A Primer*. Massachusetts: Cutter Information, 1991.
- Cummins, R. and Pollock J. eds. *Philosophy and AI: Essays at the Interface*. Massachusetts: MIT Press, 1991.
- Cymrot, A. Gelber A. and Cole L. L. "How to Invest in Office Building in Shopping Centers." *Real Estate Review* 12, no. Summer (1982): 75-82.
- Dahlén, M. and Lange F. "Real Consumer in the Virtual Store." *Scandinavian Journal of Management* 18 (2002): 341-363.
- Darian, J. C. "In-Home Shopping: Are There Consumer Segments?" *Journal of Retailing* 63, no. 2 (1987): 163-186.
- Dasgupta, C. G. Dispensa G. S. and Ghose S. "Comparing the Predictive Performance of a Neural Network Model With Some Traditional Market Response Models." *International Journal of Forecasting* 10 (1994): 235-244.

- Davies, F. Moutinho L. and Curry B. "ATM User Attitudes: a Neural Network Analysis." *Marketing Intelligence & Planning* 14, no. 2 (1996): 26-32.
- Davies, J. Foxall G. R. and Pallister J. "Beyond the Intention-Behaviour Mythology: An Integrated Model of Recycling." *Marketing Theory* 2, no. 1 (2002): 29-113.
- Dawkins, P. M. and Reichheld F. F. "Customer Retention as a Competitive Weapon." *Directors & Board* Summer (1990): 42-49.
- Dawson, J. "Key Success Factors of European Retailers." Prepared as background material for ESADE Retail marketing and Management unit 2005, 63.
- Day, R. I. and Bodur M. "Consumer Response to Dissatisfaction with Services and Intangible Products." In *Consumer Research, volume 5*, H.K. Hunt, ed., Association for consumer research, Chicago, 1978, 263-272.
- de Kare-Silver, M. *E-Shock*. New York: McMillan Business, 1998.
- de Ville, B. "Intelligent Tools for Marketing Research: Case-Based Reasoning." *Marketing Research* 9, no. 2 (1997): 38-40.
- Decker, R. and Monien K. "Market Basket Analysis with Neural Gas Networks and Self-Organizing Maps." *Journal of Targeting, Measurement and Analysis for Marketing* 11, no. 4 (2003): 373-386.
- Degeratu, A. M. Rangaswamy A. and Wu J. "Consumer Choice Behavior in Online and Traditional Supermarkets: The Effects of Brand Name, Price, and Other Search Attributes." *International Journal of Research in Marketing* 17 (2000): 55-78.
- Denzin, N. K. *Handbook of Qualitative Research*. California: Sage, 2000.
- Desmet, P. "Buying Behavior Study with Basket Analysis: Pre-Clustering with a Kohonen Map." *European Journal of Economic and Social Systems* 15, no. 2 (2001): 17-30.
- DeSouza, G. "Designing a Customer Retention Plan." *The Journal of Business Strategy* March/April (1992): 24-32.
- Devletoglou, N. E. *Consumer Behaviour: an Experiment in Analytical Economics*. United Kingdom: Harper and Row, 1971.
- Dholakia, R. R. Pedersen B. and Hikmet N. "Married Males and Shopping: Are They Sleeping Partners?" *International Journal of Retail & Distribution Management* 23, no. 3 (1995): 27-33.

- Dholakia, V. and Bagozzi R. P. *Consumer Behavior in Digital Environments*. New York: Wiley & Sons, 2001.
- Dibb, S. and Simkin L. "Targeting, Segments and Positioning." *International Journal of Retail & Distribution Management* 19, no. 3 (1991): 4-10.
- Dietterich, T. G. "Machine Learning Research: Four Current Directions." *AI Magazine* 18 (1997): 97-136.
- Dimitrov, V. "Introduction to Fuzziology." In *Fuzzy Logic: A Paradigm for the New Millenium*, V. Dimitrov and V. Korotkich, eds., Heidelberg, N.Y, 2002, 1-14.
- Dimitrov, V. and Wechler W. "Optimal Fuzzy Control of Humanistic Systems." *Proceedings of the Sixth World Congress of IFAC*, Pittsburgh, PA: Instrument Society of America, 1975.
- Dimitrov, V. and Korotkich V. *Fuzzy Logic: A paradigm for the New Millennium*. Heidelberg: Physica-Verlag, 2002.
- Douthu, N. and Garcia A. "The Internet Shopper." *Journal of Advertising Research* 39, no. 3 (1999): 52-58.
- Downs, A. "A Theory of Consumer Efficiency." *Journal of Retailing* 37, no. 1 (1961): 6-12; 50-51.
- Drucker, P. F. "The Economy's Power Shift." *The Wall Street Journal* 24, no. September (1992): A16.
- Duan, Y. and Burrell P. "A Hybrid System for Strategic Marketing Planning." *Marketing Intelligence and Planning* 13, no. 11 (1995): 5-12.
- Dubois, B. *Comprendre Le Consommateur*. France: Dalloz, 1994.
- Dubois, D. and Prade H. "What Does Fuzzy Logic Bring to AI?" *ACM Computing Surveys* 27, no. 3 (1995): 328-330.
- Duncan, G. Gorr W. and Szczypula J. "Forecasting Analogous Time Series." In *Principles of Forecasting: Handbook for Researchers and Practitioners*, J.S. Armstrong, ed., Kluwer Academic Publishers, Norwell: 2001, 195-213.
- Dunnet, A. J. and Arnold S. "Falling Prices, Rising Morale: The Impact of Organizational Culture on Customer Satisfaction at Wal-Mart." *Doctoral Colloquium, the Marketing Academy* (1999).
- Dutta, S. Shekhar S. and Wong W. Y. "Decision Support in Conservative Domains: Generalization with Neural Networks." *Decision Support Systems* 11 (1994): 527-544.

Encyclopaedia Britannica. "Learning." Web page, 2004 [accessed November 2004]. Available at <http://www.britannica.com/ebc/article?tocId=9369902&query=learning&ct=>.

East, R. *Consumer Behavior: Advances and Applications in Marketing*. United Kingdom: Prentice Hall, 1997.

East, R. Harris P. Willson G. and Lomax W. "Loyalty to Supermarkets." *International Review of Retail, Distribution and Consumer Research* 5, no. 1 (1995): 99-109.

Engel, J. F. Blackwell R. D. and Miniard P. W. *Consumer Behavior*. London: Dryden Press, 1995.

Engel, J. F. Kollat D. T. and Blackwell R. D. *Consumer Behavior*. New York: Holt, Rinehart & Winston, 1968.

Eroglu S. A. Machleit K. A. and Davis L. M. "Atmospheric Qualities of Online Retailing. A Conceptual Model and Implications." *Journal of Business Research* 54 (2001): 177-184.

Everitt, B. S. *Cluster Analysis*. London: Halsted Press, 1993.

Fazio, R. H. and Zanna M. P. "Direct Experience and Attitude-Behavior Consistency." *Advances in Experimental Social Psychology* 14 (1981): 161-202.

Feigenbaum, E. A. and Feldman J. eds. *Computers and Thought*. Massachusetts: MIT Press, 1995.

Fischler, M. A. and Firschein O. *Intelligence: The Eye, the Brain and the Computer*. Massachusetts: Addison-Wesley, 1987.

Fish, K. E. Barnes J. H. and Aiken M. W. "Artificial Neural Networks - A New Methodology for Industrial Market Segmentation." *Industrial Marketing Management* 24 (1995): 431-438.

Fish, K. E. Johnson J. D. Forsey R. E. and Blodgett J. G. "Using an Artificial Neural Network Trained With a Genetic Algorithm to Model Brand Share." *Journal of Business Research* 57 (2004): 79-85.

Fishbein, M. "An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward the Object." *Human Relations* 16 (1963): 233-40.

Fishbein, M. and Ajzen I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Massachusetts: Addison-Wesley, 1975.

Flach, P. A. "On the State of the Art in Machine Learning: A Personal Review." *Artificial Intelligence* 131 (2001): 199-222.

- Flick, U. *An Introduction to Qualitative Research*. United Kingdom: Sage, 1998.
- Foxall, G. R. "The Consumer Situation: an Integrative Model for Research in Marketing." *Journal of Marketing Management* 8 (1992): 383-404.
- Foxall, G. R. "Foundations of Consumer Behaviour Analysis." *Marketing Theory* 1, no. 2 (2001): 165-99.
- Foxall, G. R. Goldsmith R. E. and Brown S. *Consumer Psychology for Marketing*. United Kingdom: International Thomson Business Press, 1998.
- Fram, E. H. and Axelrod J. "The Distressed Shopper." *American Demographics* October (1990): 44-45.
- Frees, E. W. and Miller T. W. "Sales Forecasting Using Longitudinal Data Models." *International Journal of Forecasting* 20, no. 1 (2004): 99-114.
- Gardner, H. *Las Cinco Mentes Del Futuro: Un Ensayo Educativo*. Spain: Paidós, 2005.
- Gehrt, K. C. and Yale L. J. "The Dimensionality of the Convenience Phenomenon: A Qualitative Re-Examination." *Journal of Business and Psychology* 8, no. 2 (1993): 163-180.
- Georgoff, D. A. and Murdick R. G. "Manager's Guide to Forecasting." *Harvard Business Review* 64, no. Jan-Feb (1986): 110-120.
- Glaser, B. and Strauss A. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine, 1967.
- Godwin, P. and Wright G. "Improving Judgmental Time Series Forecasting: A Review of the Guidance Provided by Research." *International Journal of Forecasting* 9 (1993): 147-161.
- Granger, C. W. J. *Forecasting in Business and Economics*. California: Academic Press, 1999.
- Greco, A. J. and Fields D. M. "Profiling Early Tiers of Service Innovations: A Look at Interactive Home Video Ordering Services." *Journal of Services Marketing* 5, no. 3 (1991): 19-26.
- Griffith, D. A. and Krampf R. F. "Emerging Trends in US Retailing." *Long Range Planning* 30, no. 6 (1997): 847-852.
- Grossberg, S. "Adaptive Pattern Classification and Universal Recoding: I. Parallel Development and Coding of Neural Feature Detectors." *Biological Cybernetics* 23 (1976): 121-134.

- Grove, S. and Fisk R. P. "Observational Data Collection Methods for Services Marketing: An Overview." *Journal of the Academy of Marketing Science* 20, no. 3 (1992): 217-224.
- Gupta, A. Su B-c. and Walter Z. "Risk Profile and Consumer Shopping Behavior in Electronic and Traditional Channels." *Decision Support Systems* 38 (2004): 347-367.
- Gwinner, K. P. Dwayne D. G. and Bitner M. J. "Relational Benefits in Services Industries: The Customer's Perspective." *Journal of the Academy of Marketing Science* 26, no. 2 (1998): 101-115.
- Ha, S. H. and Park S. C. "Application of Data Mining Tools to Hotel Data Mart on the Intranet for Database Marketing." *Expert Systems with Applications* 15 (1998): 1-31.
- Hair, J. F. Bush J. R. P. and Ortinau D. J. *Marketing Research*. New York: McGraw-Hill Companies, 2003.
- Hammersley, M. and Atkinson P. *Ethnography Principles in Practice*. London: Tavistock, 1983.
- Hamuro, Y. Kawata H. Katoh N. and Katsutoshi Y. *A Machine Learning Algorithm for Analysing String Patterns Helps to Discover Simple and Interpretable Business Uses from Purchase History*. Berlin: Springer-Verlag, 2001.
- Hansen, T. M. Jensen J. M. and Solgaard H. S. "Predicting Online Grocery Buying Intention: a Comparison of the Theory of Reasoned Action and the Theory of Planned Behavior." *International Journal of Information Management* 24 (2004): 539-550.
- Hanson, S. J. and Burr D. J. "What Connectionist Models Learn: Learning and Representations in Connectionist Networks." *Behavioural and Brain Sciences* 13 (1990): 471-518.
- Harmon, P. and King D. *Expert Systems: Artificial Intelligence in Business*. New York: John Wiley & Sons, 1985.
- Harrell, G. *Consumer Behavior*. New York: Harcourt Brace Jovanovich, 1986.
- Harrison, T. "Editorial: Why trust is important in customer relationships and how to achieve it." *Journal of Financial Services Marketing* 7, no.3 (2003): 206-209.
- Harrison, T. and Ansell J. "Customer Retention in the Insurance Industry: Using Survival Analysis to Predict cross-selling opportunities." *Journal of Financial Services Marketing* 6, no.3 (2002): 229-239.
- Haugeland, J. ed. *Mind Design II: Philosophy, Psychology, Artificial Intelligence*. Massachusetts: MIT Press, 1997.

Hebb, D. O. "Drives and the C.N.S. (Conceptual Nervous System)." *Psychological Review* 62 (1955): 243-254.

———. *The Organization of Behavior*. New York: John Wiley & Sons, 1949.

Hendler, J. "Agents and the Semantic Web." *IEEE Intelligent Systems* 16, no. 2 (2001): 30-37.

Hergenhahn, B. R. and Olson M. *An Introduction to Theories of Learning*. New Jersey: Prentice-Hall, 1997.

Heshmaty, B. and Kandel A. "Fuzzy Linear Regression and Its Applications to Forecasting in Uncertain Environment." *Fuzzy Sets and Systems* 15 (1985): 159-191.

Hibbert, B. and Wilkinson J. F. "Chaos Theory and the Dynamics of Marketing Systems." *Journal of the Academy of Marketing Science* 22, no. 3 (1994): 218-233.

Hill, T. and Remus W. "Neural Network Models for Intelligent Support of Managerial Decision Making." *Decision Support Systems* 11 (1994): 449-459.

Hill, T. Marquez L. O'Connor M. and Remus W. "Artificial Neural Network Models for Forecasting and Decision Making." *International Journal of Forecasting* 10 (1994): 5-15.

Hintzman, D. "Human Learning and Memory." *Annual Review of Psychology* 41 (1991): 109-139.

Holbrook, M. B. and Hirschman E. C. "The Experiential Aspects of Consumption: Consumer Fantasies, Feelings and Fun." *Journal of Consumer Research* 9, no. September (1982): 132-140.

Hopfield, J. J. "Neural Networks and Physical Systems with Emergent Collective Computational Abilities." *In Proceedings of the National Academy of Sciences, National Academy of Sciences*. (1982): 2554-2558.

Howard, J. A. *Buyer Behavior in Marketing Strategy*. New Jersey: Prentice-Hall, 1994.

Howard, J. A. and Sheth J. N. *The Theory of Buyer Behavior*. New York: John Wiley & Sons, 1969.

Howell, R. and Rogers J. "Research into Shopping Mall Choice Behaviour." *Advances in Consumer Research* (1980): 671-687.

———. "Market Definition and Segmentation Using Fuzzy Clustering Methods." *International Journal of Research in Marketing* 3, no. 2 (1986): 117-134.

- Huang, W. Y. and Lippman R. P. "Comparisons between Neural Net and Conventional Classifiers." *IEEE 1st International Conference on Neural Networks IV* (1997): 485-493.
- Huth, L. W. Epright R. D. and Taube M. P. "The Indexes of Consumer Sentiment and Confidence: Leading or Misleading Guides to Future Buyer Behavior." *Journal of Business Research* 29, (1994): 199-206.
- Häubl, G. and Trifts V. "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Aids." *Marketing Science* 19, no. 1 (2000): 4-21.
- Infosino, J. W. "Forecasting New Products Sales from Likelihood of Purchase Rating." *Marketing Science* 5, no. Fall (1986): 372-390.
- Jackson, P. *Introduction to Expert Systems*. United Kingdom: Addison-Wesley, 1990.
- Jain, A. K. Murty M. N. and Flynn P. J. "Data Clustering: A Review." *ACM Computing Surveys* 31 (1999): 264-323.
- Jamieson, L. F. and Bass F. M. "Adjusting Stated Intention Measures To Predict Trial Purchase of New Products: A Comparison of Models and Methods." *Journal of Marketing Research* 26, no. August (1989): 336-345.
- Jones, M. A. Mothersbaugh D. L. and Beatty S. E. "Switching Barriers and Repurchase Intentions in Services." *Journal of Retailing* 76, no. 2 (2000): 259-274.
- Jurik, M. "Consumers' Guide to Neural Network Software." *Futures: The Magazine of Commodities and Options* (1993): 36-41.
- Juster, F. T. "Consumer Buying Intentions and Purchase Probability: an Experiment in Survey Design." *Journal of the American Statistical Association* 61, no. September (1966): 658-696.
- Kahn, B.E. and Schmittlein D. C. "Shopping Trip Behavior: an Empirical Investigation." *Marketing Letters* 1, no. 1 (1989): 55-69.
- Kalakota, R. and Whinston A. B. *Electronic Commerce, A Manager's Guide*. Massachusetts: Addison Wesley, 1997.
- Kalwani, M. U. and Silk A. J. "On the Reliability and Predictive Validity of Purchase Intention Measures." *Marketing Science* 1, no. Summer (1982): 243-286.
- Kamakura, W. A. and Gessner G. "Consumer Sentiment and Buying Intentions Revisited: a Comparison of Predictive Usefulness." *Journal of Economic Psychology* 7, no. 2 (1986): 197-221.

Kaufman-Scarborough, C. "A New Look at One Stop Shopping: a TIMES Model Approach to Matching Store Hours and Shopping Schedules." *Journal of Consumer Marketing* 13, no. 1 (1996): 4-25.

Kaufman-Scarborough, C. and Lindquist J. D. "E-Shopping in a Multiple Channel Environment." *Journal of Consumer Marketing* 19, no. 4 (2002): 333-350.

———. "Understanding the Experience of Time Scarcity: Linking Consumer Time-Personality and Marketplace Behavior." *Time & Society* 12, no. 2 (2003): 349-370.

Keaveney, S. M. "Customer Switching Behaviour in Services Industries: An Exploratory Study." *Journal of Marketing* 59, no. April (1995): 71-82.

Keaveney, S. M. and Parthasarathy M. "Customer Switching Behavior in Online Services: An Exploratory Study of the Role of Selected Attitudinal, Behavioral, and Demographic Factors." *Journal of the Academy of Marketing Science* 29, no. 4 (2001): 374-390.

Kim, B-D. and Park K. "Studying Patterns of Consumer's Grocery Shopping Trip." *Journal of Retailing* 73, no. 4 (1997): 501-517.

Kim, E. Y. and Kim Y-K. "Predicting Online Purchase Intentions for Clothing Products." *European Journal of Marketing* 38, no. 7 (2004): 883-897.

Kim H-S. and Cho S-B. "Application of Interactive Genetic Algorithm to Fashion Design." *Engineering Applications of Artificial Intelligence* 13 (2000): 635-644.

Kinney, T. C. and Taylor J. R. *Marketing Research: An Applied Approach*. New York: McGraw-Hill, 1996.

Klapper, D. and Herwartz H. "Forecasting Market Share Using Predicted Values of Competitive Behavior: Further Empirical Results." *International Journal of Forecasting* 16, no. 3 (2000): 399-421.

Klein, L. R. and A. S. Goldberger. *An Econometric Model of the United States 1929-1952*. Amsterdam: North-Holland, 1955.

Klein, L. R. and Young R. M. *An Introduction to Econometric Forecasting and Forecasting Models*. Kentucky: D.C. Heath and Company, 1980.

Korf, D. E. "Does Deep Blue Use Artificial Intelligence?" *ICCA Journal* 20, no. 4 (1997).

Kosko, B. *Pensamiento Borroso. La Nueva Ciencia De La Lógica Borrosa*. Spain: Crítica, 1995. Koslowsky, S. "Can You Predict the Future?" *Targetmarketingmag.com* October (2003): 87-90.

- Kotler, P. "Atmosphere as a Marketing Tool." *Journal of Retailing* 49, Winter (1973): 48-64.
- Kotler, P. *Marketing Management*. New Jersey: Prentice-Hall, 2000.
- Kotler, P. and Armstrong G. *Principles of Marketing*. New Jersey: Prentice Hall, 1994.
- Krycha, K. A. and Wagner U. "Applications of Artificial Neural Networks in Management Science: A Survey." *Journal of Retailing and Consumer Services* 6 (1999): 185-203.
- Kumar, A. Rao V. R. and Soni H. "An Empirical Comparison of Neural Network and Logistic Regression Models." *Marketing Letters* 6, no. 4 (1995): 251-263.
- Kuo, R. J. Wu P. and Wang C. P. "Sales Forecasting System through Integration of Artificial Neural Networks and Fuzzy Networks with Fuzzy Weight Elimination." *Neural Networks* 15, no. September (2002): 909-925.
- LaBarbera, P. A. and Mazursky D. "A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction: The Dynamic Aspect of the Cognitive Process." *Journal of Marketing Research* 20, no. November (1983): 393-404.
- Lariviere, B. and Van Den Poel D. "Investigating the Role of Product Features in Preventing Customer Churn, by Using Survival Analysis and Choice Modelling: The Case of Financial Service." *Working Papers of Faculty of Economics and Business Administration*, Ghent University, Belgium (2004).
- Latacz-Lohmann, U. and Foster C. "From "Niche" to "Mainstream" - Strategies for Marketing Organic Food in Germany and the UK." *British Food Journal* 99, no. 8 (1997): 275-282.
- Lavin, M. "Wife's Employment, Time Pressure, Phone and Mail Order Shopping- an Exploratory Study." *Journal of Direct Marketing* 7, no. 1 (1993): 42-49.
- Leake, D. "Artificial Intelligence." In *Van Nostrand's Scientific Encyclopaedia. Two-Volume Set*, G.D. Considine, ed., John Wiley & Sons, New York, 2002.
- Lee, D. H. Kim S. H. and Ahn B. S. "A Conjoint Model for Internet Shopping Malls Using Customer's Purchasing Data." *Expert Systems with Applications* 19 (2000): 59-66.
- Lee, M-S. Elango B. and Schnaars S. P. "The Accuracy of the Conference Board's Buying Plans Index: A Comparison of Judgmental Vs. Extrapolation Forecasting Methods." *International Journal of Forecasting* 13 (1997): 127-135.

- Leeflang, P. S. H. and van Raaij W. F. "The Changing Consumer in the European Union: a "Meta Analysis"." *International Journal of Research in Marketing* 12 (1995): 373-383.
- Lemon, K. N. White T. B. and Winer R. S. "Dynamic Customer Relationship Management: Incorporating Future Considerations into the Service Retention Decision." *Journal of Marketing* 66, no. 1 (2002): 1-14.
- Lempert, P. *Being the Shopper. Understanding the Buyer's Choice*. New York: John Wiley & Sons, 2002.
- Li, S. "The Development of a Hybrid Intelligent System for Developing Marketing Strategy." *Decision Support Systems* 27 (2000): 395-409.
- Lilien, G. L. and Kotler P. *Marketing Decision-Making: A Model-Building Approach*. New York: Harper & Row, 1983.
- Loudon, D. L. and Della Vitta A. J. *Consumer Behavior: Concepts and Applications*. Singapore: McGraw-Hill, 1993.
- Luger, F. G. *Artificial Intelligence, Structures and Strategies for Complex Problem Solving*. Massachusetts: Addison Wesley, 2002.
- Madrigal, R. "Social Identity Effects in a Belief-Attitude-Intentions Hierarchy: Implications for Corporate Sponsorship." *Psychology and Marketing* 18, no. 2 (2001): 145-165.
- Makridakis, S. "The Value of Empirical Work: A Personal View." *International Journal of Forecasting* 9 (1993): 28-29.
- Makridakis, S. and Wheelwright S. C. *Forecasting: Methods and Applications*. New York: John Wiley & Sons, 1978.
- . *Forecasting Methods for Management*. Singapore: John Wiley & Sons, 1989.
- Makridakis, S. Wheelwright S. C. and Hyndman R. J. *Forecasting: Methods and Applications*. New York: John Wiley & Sons, 1998.
- Malhotra, N. K. and Birks D. F. *Marketing Research: An Applied Approach*. United Kingdom: Prentice Hall, 2000.
- Malhotra, N. K. Peterson M. and Kleiser S. B. "Marketing Research: A State-of-the-Art. Review and Directions for the Twenty-First Century." *Journal of the Academy of Marketing Science* 27, no. 2 (1999): 160-183.

- Mandani, E. H. "Applications of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis." *Proceedings of the IEEE (Trans. Computers)* 126 (1977): 1182-1191.
- Manski, C. F. "The Use of Intentions Data to Predict Behavior: A Best Case Analysis." *Journal of the American Statistical Association* 85, no. December (1990): 934-940.
- Mansur, Y. *Fuzzy Sets and Economics: Applications of Fuzzy Sets to Non Cooperative Oligopoly*. United Kingdom: Edward Elgar, 1995.
- Marr, D. "AI: a Personal View." *Artificial Intelligence* 9, no. 1 (1977): 129-142.
- Maslow, A. H. *Motivation and Personality*. New York: Harper, 1954.
- Mathwick, C. Malhotra N. K. and Rigdon E. "The Effect of Dynamic Retail Experiences on Experiential Perceptions of Value: an Internet and Catalogue Comparison." *Journal of Retailing* 78, no. 1 (2002): 51-60.
- Mazanec, J. A. "Classifying Tourists into Market Segments: A Neural Network Approach." *Journal of Travel & Tourism Marketing* 1, no. 1 (1992): 39-59.
- McCarthy, J. *Formalizing Common Sense: Papers by John McCarthy*. New Jersey: Ablex Publishing Corporation, 1990.
- McCulloch, W. S. and Pitts W. H. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics* 5 (1943): 115-137.
- McGoldrick, P. *Retail Marketing*. United Kingdom: McGraw-Hill, 2002.
- McGoldrick, P. J. and Andre E. "Consumer Misbehaviour: Promiscuity or Loyalty in Grocery Shopping." *Journal of Retailing and Consumer Services* 4, no. 2 (1997): 73-81.
- McNeil, J. M. "Federal Programs to Measure Consumer Purchase Expectations, 1946-73: A Post Mortem." *Journal of Consumer Research* 1, no. December (1974): 1-10.
- McQuarrie, E. F. "An Alternative to Purchase Intentions: the Role of Prior Behavior in Consumer Expenditure on Computers." *Journal of the Market Research Society* 30, no. October (1988): 407-437.
- Mentzer, J. T. Bienstock C. C. and Kahn K. B. "Benchmarking Sales Forecasting Management." *Business Horizons* 42 (1999): 48-56.
- Mercer, D. *Marketing Strategy: The Challenge of the External Environment*. California: Sage, 1998.

Miller, G. A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *The Psychological Review* 63, no. March (1956): 81-97.

Minsky, M. and Papert S. A. *Perceptrons: An Introduction to Computational Geometry*. Massachussets: MIT Press, 1969.

Mir, P. González J. and Gil A. "Ús de les tecnologies de la informació i les comunicacions a la ciutat de Lleida." Web page, January 2002 [accessed December 2003]. Available at <http://www.lapaeria.es>.

MIT Encyclopedia of Cognitive Science. "Machine Learning." Web page, [accessed December 2004a]. Available at <http://cognet.mit.edu/library/erefs/mitecs/dietterich.html>.

MIT Encyclopedia of Cognitive Science. "Unsupervised Learning." Web page,[accessed December 2004]. Available at <http://cognet.mit.edu/library/erefs/mitecs/dayan.html>.

Mitchell, A. A. Russo J. E. and Wittink D. R. "Issues in the Development and Use of Expert Systems for Marketing Decisions." *International Journal of Research in Marketing* 8 (1991): 41-50.

Mitchell, T. M. *An Introduction to Genetic Algorithms*. Massachusetts: MIT Press, 1996.

Mittal, B. and Kamakura W. A. "Satisfaction, Repurchase Intent and Repurchase Behaviour: Investigations the Moderation Effect of Customer Characteristics." *Journal of Marketing Research* 38, no. February (2001): 131-142.

Mittal, B. and Lassar W. M. "Why Do Customers Switch? The Dynamics of Satisfaction Versus Loyalty." *The Journal of Services Marketing* 12, no. 3 (1998): 177-194.

Moe, W. and Fader P. S. "Capturing Evolving Visit Behaviour in Click stream Data." Wharton Marketing Department 2001.

———. "Dynamic Conversion Behaviour at E-Commerce Sites." Wharton Marketing Department 2002.

Montgomery, A. L. "Applying Quantitative Marketing Techniques to the Internet." *Interfaces* 31, no. 2 (2001): 91-108.

Montgomery, D. Swinnen G. and Vanhoof K. "Comparison of Some AI and Statistical Classification Methods for a Marketing Case." *European Journal of Operational Research* 103 (1997): 312-325.

Morrison, D. G. "Purchase Intentions and Purchase Behavior." *Journal of Marketing* 43, no. Spring (1979): 65-74.

- Morwitz, V. G. and Schmittlein D. "Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which 'Intenders' Actually Buy?" *Journal of Marketing Research* 29, no. November (1992): 391-405.
- Morwitz, V. G. Steckel J. and Gupta A. "When Do Purchase Intentions Predict Sales?" *Marketing Science Institute*, Working paper (1997): 97-112.
- Mowen, J. C. *Consumer Behavior*. New Jersey. Prentice-Hall, 1995.
- Moye, L. N. and Kincade D. H. "Influence of Usage Situations and Consumer Shopping Orientations on the Importance of the Retail Store Environment." *The International Review of Retail, Distribution and Consumer Research* 12, no. 1 (2002): 59-79.
- Mozer, M. C. Wolniewicz R. Grimes D. B. Johnson E. Kaushansky H. "Predicting Subscriber Dissatisfaction and Improving Retention Inthe Wireless Telecommunications Industry." *IEEE Transactions on Neural Networks* 11, no. 3 (2000): 690-696.
- Mulhern, F. J. "Retail Marketing: From Distribution to Integration." *International Journal of Research in Marketing* 14 (1997): 103-124.
- Munakata, T. and Jani Y. "Fuzzy Systems: An Overview." *Communications of the ACM* 37 (1994): 69-76.
- Myers, D. G. *Psychology*. New York: Worth Publishers, 2001.
- Newell, A. "Some Intellectual Issues in the History of Artificial Intelligence." In *The Study of Information: Interdisciplinary Perspectives*, F. Machlup and U. Mansfield, eds., John Wiley & Sons, 1993, 187-227.
- Newell, A. and Simon H. *Human Problem Solving*. New Jersey: Prentice-Hall, 1972.
- Newman, A. J. and Cullen P. *Retailing: Environment and Operations*. United Kingdom: Thomson Learning International, 2002.
- Newman, A. J. Yu D. K. C. and Oulton D. P. "New Insights into Retail Space and Format Planning From Customer-Tracking Data." *Journal of Retailing and Consumer Services* 9 (2002): 253-258.
- Nicosia, F. *Consumer Decision Processes: Marketing and Advertising Implications*. New Jersey: Prentice-Hall, 1966.
- O'shaughnessy, J. *Explaining Buyer Behaviour: Central Concepts and Philosophy of Social Science*. New York: Oxford University Press, 1992.
- Oliver, R. L. "Whence Consumer Loyalty." *Journal of Marketing* 63, no. Special issue (1999): 33-44.

- Oppewal, H. Louvieve J. and Timmermans H. "Modelling Hierarchical Conjoint Processes with Integrated Choice Experiments." *Journal of Marketing Research* 31 (1994): 92-105.
- Ozer, M. "User Segmentation of Online Music Services Using Fuzzy Clustering." *Omega: The International Journal of Management Science* 29, no. 2 (2001): 193-206.
- Padmanabham, B. Zheng Z. and Kimbrough S. O. "Personalization From Incomplete Data: What You Don't Know Can Hurt." in *Proceeding of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2001).
- Paliouras, G. Papatheodorou C. Karkaletsis V. and Spyropoulos C. D. "Discovering User Communities on the Internet Using Unsupervised Machine Learning Techniques." *Interacting With Computers* 14 (2002): 761-791.
- Parasuraman, A. and Colby C. L. *Techno-Ready Marketing: How and Why Your Customers Adopt Technology*. New York: Free Press, 2001.
- Partridge, D. and Wilks Y. A. *The Foundations of Artificial Intelligence, a Sourcebook*. United Kingdom: Cambridge University Press, 1990.
- Pavlov, I. P. *Conditioned Reflexes*. United Kingdom: Oxford University Press, 1927.
- Piera, N. and Aguilar-Martin, J. "Controlling Selectivity in Non-Standard Pattern Recognition Algorithms." *IEEE Transactions on SMC volume* 21, no. 1 (1991).
- Popkowski, L. P. T. L. Sinha A. and Timmermans H. J. P. "Consumer Store Choice Dynamics: An Analysis of the Competitive Market Structure for Grocery Stores." *Journal of Retailing* 76, no. 3 (2000): 323-345.
- Raijas, A. "The Consumer Benefits and Problems in the Electronic Grocery Store." *Journal of Retailing and Consumer Services* 9 (2002): 107-113.
- Ranaweera, C. and Prabhu J. "The Influence of Satisfaction, Trust and Switching Barriers on Customer Retention in a Continuous Purchasing Setting." *International Journal of Service Industry Management* 14, no. 4 (2003): 374-395.
- Ray, A. "How to Encourage Internet Shopping." *Marketing* May, no. 3 (2001): 41-42.
- Reibstein, D. J. "What Attracts Customers to Online Stores, and What Keeps Them Coming Back?" *Journal of the Academy of Marketing Science* 30, no. 4 (2002): 465-473.
- Reichheld, F. F. "Learning From Customer Defections." *Harvard Business Review* 74, no. 2 (1996b): 56-69.

———. *The Loyalty Effect: The Hidden Force behind Growth, Profits and Lasting Value*. Massachusetts: Harvard Business School Press, 1996a.

Reichheld, F. F. and Sasser W. E. "Zero Defections: Quality Comes to Services." *Harvard Business Review* 68, no. 5 (1990): 105-111. Reimers, V. and Clulow V. "Shopping and Convenience: A Model for Retail Centres." *ANZMAC Conference*, 2000.

Retail I.T. "23rd Annual Survey of Retail Information Technology." *Chain Store Age* November, no. 2 (2004): 4A-27A.

Rich, E. and Knight K. *Artificial Intelligence*. Indiana: McGraw-Hill, 1991.

Ronis, D. L. Yates J. F. and Kirscht J. P. "Attitudes, Decisions and Habits as Determinants of Repeated Behavior." In *Attitude Structure and Function*, R. Anthony S. Pratkanis and J. Greckler, eds., Erlbaum, Hillsdale, N.J, 1989, 213-239.

Roos, I. "Switching Processes in Customer Relationships." *Journal of Service Research* 2, no. 1 (1999): 68-85.

Roos, I. Edvardsson B. and Gustafsson A. "Customer Switching Patterns in Competitive and Non-competitive Service Industries." *Journal of Service Research* 6, no. February (2004): 256-271.

Rosenberg, L. G. and Czepiel J. A. "A Marketing Approach for Customer Retention." *Journal of Consumer Marketing* 1 (1984): 45-51.

Rosenberg, M. J. and Hovland C. I. "Cognitive, Affective and Behavioral Components of Attitudes." In *Attitude organization and change*, C.I. Hovland and M.J. Rosenberg, eds., Yale University Press, New Haven, 1960, 1-14.

Rosenblatt, F. "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review* 65 (1958): 386-407.

Rumelhart, D. E. Hinton G. E. and Williams R. J. "Learning Representations by Back-Propagating Errors." *Nature* 323 (1986): 533-36.

Russell, S. J. and Norvig P. *Artificial Intelligence: a Modern Approach*. New Jersey: Prentice-Hall International, 1995.

Rust, R. T. and Zahorik A. J. "Customer Satisfaction, Customer Retention, and Market Share." *Journal of Retailing* 69, no. 2 (1993): 193-215.

Samli, A. C. *Strategic Marketing for Success in Retailing*. Connecticut: Quorum Books, 1998.

Schmittlein, D. C. and Peterson R. A. "Customer Base Analysis: An Industrial Purchase Process Application." *Marketing Science* 13, no. 1 (1994): 41-67.

Searle, J. R. "Minds, Brains and Programs." *Behavioral and Brain Sciences* 3, no. 3 (1980): 417-424.

Searle, J. R. *Minds, Brains and Science*. Massachusetts: Harvard University Press, 1984.

Sejnowski, T. J. and Rosenberg C. R. *NETtalk: a Parallel Network That Learns to Read Aloud*. John Hopkins University, Baltimore, MD: Technical Report JHU/EECS-86/01, 1986.

Serrano-Cinca, C. "Self Organizing Neural Networks for Financial Diagnosis." *Decision Support Systems* 17, no. July (1996): 227-238.

Setiono, R. Thong J. Y. L. and Yap C. S. "Symbolic Rule Extraction from Neural Networks - An Application to Identifying Organizations Adopting IT." *Information & Management* 34, no. 2 (1998): 91-101.

Shanker, M. Hu M. Y. and Hung M. S. "Effect of Data Standardization on Neural Network Training." *Omega: The International Journal of Management Science* 24, no. 4 (1996): 385-97.

Sheth, J. N. and Mittal B. *Customer Behavior: A Managerial Perspective*. Ohio: Thomson- South Western, 2004.

Shim, S. and Drake M. F. "Consumer Intention to Utilize Electronic Shopping." *Journal of Direct Marketing* 4, no. 3 (1990): 22-33.

Shim, S. Eastlick M. A. and Lotz S. L. "Assessing the Impact of Internet Shopping on Store Shopping Among Mall Shoppers and Internet Users." *Journal of Shopping Centre Research* 7, no. 2 (2000): 7-43.

Silk, A. J. and Urban G. L. "Pre-Test Market Evaluation of New Packaged Goods: A Model and Measurement Methodology." *Journal of Marketing Research* January (1978): 1-29.

Simon, H. "Artificial Intelligence: An Empirical Science." *Artificial Intelligence* 77, no. 1 (1995): 95-127.

Skinner, B. F. *Science and Human Behavior*. New York: Macmillan, 1953.

Sloman, A. "AI and the study of mind." Web page, October 2002 [accessed October 2003]. Available at www.cs.bham.ac.uk/axs/misc/talks/.

———. *The Computer Revolution in Philosophy*. United Kingdom: Harvester Press, 1978.

Smithson, M. *Fuzzy Set Analysis for the Behavioral and Social Sciences*. New York: Springer-Verlag, 1987.

Stake, R. E. "Case Studies." In *Handbook of Quantitative Research*, N.K. Denzin and Y.S. Lincoln, eds., Sage, Newbury Park, CA, 1994.

Stanford University. "Branches of AI." Web page, [accessed December 2004]. Available at <http://www-formal.stanford.edu/jmc/whatisai/node2.html>.

Steenkamp, J. B. E. M. and Wedel M. "Fuzzy Cluster wise Regression in Benefit Segmentation: Application and Validation into Its Validity." *Journal of Business Research* 26, no. 3 (1992): 237-249.

Storbacka, K. Strandvik T. and Grönroos C. "Managing Customer Relationship for Profit: The Dynamics of Relationship Quality." *International Journal of Service Industry Management* 5, no. 5 (1994): 21-38.

Strader, T. J. and Shaw M. J. "Characteristics of Electronic Markets." *Decision Support Systems* 21 (1997): 185-198.

Strauss, A. and Corbin J. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. United Kingdom: Sage, 1990.

———. *Grounded Theory in Practice*. California: Sage, 1997.

Supphellen, M. and Nysveen H. "Drivers of Intention to Revisit the Web Sites of Well.Known Companies." *International Journal of Market Research* 43, no. 3 (2001): 341-352.

Tam, K. Y. "Neural Networks for Decision Support." *Decision Support Systems* 11 (1994): 389-392.

Tam, K. Y. and Kiang G. M. Y. "Managerial Applications of the Neural Networks: The Case of Bank Failure Predictions." *Management Science* 38, no. 7 (1992): 926-947.

Tao, Y-H. and Yeh C-C. R. "Simple Database Marketing Tools in Customer Analysis Retention." *International Journal of Information Management* 23 (2003): 291-301.

Tauber, E. M. "Predictive Validity in Consumer Research." *Journal of Advertising Research* 15, no. October (1975): 59-64.

Tchamova, A. and Semerdjiev T. "Fuzzy Logic Approach to Estimating Tendencies in Target Behavior." *Information & Security, an International Journal* 9 (2002): 58-69.

- Tedesco, B. G. "Neural Marketing: Artificial Intelligence Neural Networks In Measuring Consumer Expectation." Web page 1992 [accessed April 2003]. Available at <http://www.tedescoanalytics.com/>.
- Thall, N. "Neural Forecasts: A Retail Sales Booster." *Discount Merchandiser* 32, no. 10 (1992): 41-42.
- Thomas, L. C. "A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers." *International Journal of Forecasting* 16 (2000): 14-172.
- Thorndike, E. L. *Animal Intelligence*. New York: Macmillan, 1911.
- Timmermans, H. Van der Heidjen R. and Westerveld H. "Cognition of Urban Retailing Structures: a Dutch Case Study." *Tijdschrift Voor Econ.En.Soc. Geografie* 73 (1982): 2-12.
- Tseng, F-M. and Tzeng G-H. "A Fuzzy Seasonal ARIMA Model for Forecasting." *Fuzzy Sets and Systems* 126 (2002): 367-376.
- Turing, A. M. "Computing Machinery and Intelligence." *Mind* 49 (1950): 433-460.
- Turksen, I. B. and Willson I. A. "A Fuzzy Set Model for Market Share and Preference Prediction." *European Journal of Operational Research* 82 (1995): 39-52.
- Uncles, M. D. Ehrenberg A. S. C. and Hammond K. "Patterns of Buyer Behavior: Regularities, Models and Extensions." *Marketing Science* 14, no. 3 (1995): 71-78.
- Van den Poel, D. and Buckinx W. "Predicting Online-Purchasing Behaviour." *European Journal of Operational Research* 166, no. October (2005): 557-575.
- Van Den Poel, D. and Lariviere B. "Customer Attrition Analysis for Financial Services Using Proportional Hazard Models." *European Journal of Operational Research* 157, no. 1 (2004): 196-217.
- Vellido, A. Lisboa P. J. G. and Meehan K. "Segmentation of the on-Line Shopping Market Using Neural Networks." *Expert Systems with Applications* 17 (1999): 303-314.
- Vellido, A. Lisboa P. J. G. and Vaughan J. "Neural Networks in Business: a Survey of Applications (1992-1998)." *Expert Systems with Applications* 17 (1999): 51-70.
- Venugopal V. and Baets W. "Neural Networks and Statistical Techniques in Marketing Research: A Conceptual Comparison." *Marketing Intelligence & Planning* 12, no. 7 (1994): 30-38.

Verhoef, P. C. Franses P. H. and Hoekstra J. C. "The Effect of Relational Constricts on Customer Referrals and Number of Services Purchased from a Multi service Provider: Does Age of Relationship Matter?" *Journal of the Academy of Marketing Science* 30, no. 3 (2002): 202-216.

Wang, S. and Archer N. P. "A Neural Network Technique in Modelling Multiple Criteria Multiple Person Decision Making." *Computer & Operations Research* 21, no. 2 (1994): 127-142.

Wang, S. J. Ohno-Machado L. Fraser H. S. F and Kennedy R. L. "Using Patient-Reportable Clinical History Factors to Predict Myocardial Infarction." *Computers in Biology and Medicine* 31 (2001): 1-13.

Warshaw, P. R. "Predicting Purchase and Other Behaviors from General and Contextually Specific Intentions." *Journal of Marketing Research* 17, no. February (1980): 26-33.

Watanabe, S. *Pattern Recognition: Human and Mechanical*. New York: John Wiley & Sons, 1985.

Watson, J. "Psychology as the Behaviourist Sees It." *Psychological Review* 20 (1913): 158-177.

Wedel, M. and Steenkamp J-B. E. M. "A Cluster wise Regression Method for Simultaneous Fuzzy Market Structuring and Benefit Segmentation." *Journal of Marketing Research* 28, no. November (1991): 385-396.

Weerahandi, S. and Moitra S. "Using Survey Data to Predict Adoption and Switching for Services." *Journal of Marketing Research* 32, no. 1 (1995): 85-96.

Wei, C-P. and Chiu I-T. "Tuning Telecommunications Call Detail to Churn Prediction: A Data Mining Approach." *Expert Systems with Applications* 23 (2002): 103-112.

West, P. M. Brockett P. L. and Golden L. L. "A Comparative Analysis of Neural Networks and Statistical Methods for Predicting Consumer Choice." *Marketing Science* 16, no. 4 (1997): 370-391.

Whitby, B. *Artificial Intelligence, a Handbook of Professionalism*. United Kingdom: Ellis Horwood Limited, 1988.

White, H. "Neural Network Learning and Statistics." *AI Expert* 4, no. 12 (1989): 48-82.

Wilkie, W. L. *Consumer Behavior*. New York: John Wiley & Sons, 1994.

- Willems, T. M. and Brandts L. E. M. W. "Sub-Symbolic Management Decision Support System: an Application of Neural Networks in Management." *Production Planning & Control* 8, no. 2 (1997): 123-132.
- Wilson, R. L. "A Neural Network Approach to Decision Alternative Priorization." *Decision Support Systems* 11 (1994): 431-447.
- Winston, P. H. *Artificial Intelligence*. Massachusetts: Addison-Wesley, 1992.
- Wooldridge, M. *An Introduction to Multiagent Systems*. United Kingdom: John Wiley & Sons, 2002.
- Wray, B. Palmer A. and Bejou D. "Using Neural Networks Analysis to Evaluate Buyer-Seller Relationships." *European Journal of Marketing* 28, no. 10 (1994): 32-48.
- Yakoubov, Y. Y. and Haberman S. "Review of Actuarial Applications of Fuzzy Set Theory." *Actuarial Science and Statistics Actuarial Research Report* no. 105 (1998).
- Yoon, Y. Swales G. and Margavio T. M. "A Comparison of Discriminant Analysis versus Artificial Neural Networks." *Journal of Operational Research Society* 44, no. 1 (1993): 51-60.
- Zadeh, L. A. "The Calculus of Fuzzy If-Then Rules." *AI Expert* 7 (1992): 23-27.
- . "Fuzzy Logic = Computing With Words." *IEEE Transactions on Fuzzy Systems* 4, no. 2 (1996): 103-111.
- . "Fuzzy Sets." *Proceedings of the IEEE (Information and Control)* 8, no. June (1965): 338-353.
- . "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes." *Proceedings of the IEEE (Transactions on Systems, Man and Cybernetics)* 3 (1973): 28-44.
- Zaltman, G. *How Customers Think: Essential Insights into the Mind of the Market*. Massachusetts: Harvard Business School Press, 2003.
- Zanna, M. P and Rempel J. K. "Attitudes: A New Look at an Old Concept." In *The Social Psychology of Knowledge*, D. Bar-Tal and A.W. Kruglanski, eds., Cambridge University Press, United Kingdom, 1988, 315-334.
- Zeithaml, V. A. Berry L. L. Parasuraman A. "The Behavioural Consequences of Service Quality." *Journal of Marketing* 60, no. April (1996): 31-46.

Zeithaml, V. and Gilly M. C. "Characteristics Affecting the Acceptance of Retailing Technologies: A Comparison of Elderly and Non elderly Consumers." *Journal of Retailing* 63, no. 1 (1987): 49-68.

Zhao, L. Collopy F. and Kennedy M. " The Problem of Neural Networks in Business Forecasting: An Attempt to Reproduce the Hill, O'Connor and Remus Study." *Sprouts: Working Papers on Information Environments. Systems and Organizations* 3, no. Fall (2003): 1-11.

Ziliani, C. "Retail Micro-Marketing: Strategic Advance or Gimmick?" *The International Review of Retail, Distribution and Consumer Research* 10, no. 4 (2000): 355-368.

Ziliani, C. and Bellini S. "Retail Micro-Marketing Strategies and Competition." *International Review of Retail, Distribution and Consumer Research* 14, no. January (2004): 7-18.