Journal of Marketing Management
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t914689377

Targeting direct marketing campaigns by neural networks

Gianluigi Guido[a]; M. Irene Prete[a]; Stefano Miraglia[b]; Irma De Mare[c]
[a] University of Salento, Italy [b] Imperial College London, UK [c] Leroy Merlin, Italy

First published on: 04 March 2011

To link to this Article: DOI: 10.1080/0267257X.2010.543018
URL: http://dx.doi.org/10.1080/0267257X.2010.543018

Journal of Marketing Management
i*First* 2011, 1–15

Taylor & Francis
Taylor & Francis Group

Targeting direct marketing campaigns by neural networks

Gianluigi Guido, *University of Salento, Italy*
M. Irene Prete, *University of Salento, Italy*
Stefano Miraglia, *Imperial College London, UK*
Irma De Mare, *Leroy Merlin, Italy*

Abstract Market-oriented companies increasingly aim at maximising the return of targeted direct marketing campaigns, rather than trying to reach customers and prospects indistinctly, according to a mass marketing approach. The profitability of direct marketing campaigns depends on a detailed definition of prospects and an accurate prediction of the response rate. This study shows how the use of Artificial Neural Networks (ANNs) can improve the effectiveness of direct mail marketing campaigns thanks to a better prediction of the response rate for subjects included in the target population according to factors that are believed to have an impact on their purchase intention. Results show the effectiveness of ANNs – in comparison with multiple regression analysis and logistic regression analysis – in identifying complex relationships among the data, and particularly in profiling customers and prospects and anticipating their behaviour.

Keywords direct marketing; mailing campaigns; marketing campaigns; artificial neural networks; promotions

Introduction

Market-oriented companies increasingly hold a direct marketing approach, focusing on prospects that are most likely to be enticed by particular offers, as opposed to a traditional mass marketing approach whose promotional activities are addressed to customers and prospects indistinctly. Direct marketing, defined by the Direct Marketing Association as the 'interactive system of marketing which uses one or more advertising media to effect a measurable response and/or transaction at any location', has gained momentum as companies improve the profitability of direct promotional initiatives by selectively targeting their prospects. Over the last decade, companies' expenditures in this area have considerably increased, as well as their ability to maximise profits by contacting the most promising customers (Barwise & Farley,

ISSN 0267-257X print/ISSN 1472-1376 online

DOI: 10.1080/0267257X.2010.543018
http://www.informaworld.com

2005). Furthermore, the introduction of the Internet has reduced the operational costs of direct marketing, so that even a response rate of .5% can be frequently considered adequate for a profitable e-mail marketing campaign (Direct Marketing Association, 2009).

Within a direct marketing approach, identifying best prospects – consumers who are most likely to respond to promotional activities – requires not only a deep sociodemographic analysis of customer databases, but also the development of models that point out the characteristics of those best prospects (Bhattacharyya, 1999; Kaefer, Heilman, & Ramenofsky, 2005). In the marketing field, a number of studies have comprehensively examined how consumers behave in particular settings, by developing theories and models that try to predict individual attitudes, preferences, and decisions. Specifically, several techniques have been implemented in order to improve the effectiveness of direct marketing campaigns (Bose & Chen, 2009). Database marketing is another form of direct marketing, which applies statistical analysis and informational models to individual-level data sets (Drozdenko & Drake, 2002). The advances of database technologies, the advent of the Internet and e-commerce, and the subsequent personalisation of websites through one-to-one web marketing (Pepper & Rogers, 1993) have remarkably empowered customer relationship management (CRM) (Payne & Frow, 2005).

In order to identify the factors that impact on individuals' purchase intentions and to predict the response rate of a direct marketing campaign, this study uses an innovative analytical method, the Artificial Neural Network (ANN), defined as 'a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use' (Haykin, 1999, p. 2). Results show the effectiveness of the ANN – in comparison with the two conventional statistical models of multiple regression analysis (MRA) and logistic regression analysis (LRA) – in identifying complex relationships among the data, profiling customers and prospects involved in a promotion, and anticipating their purchasing behaviour.

Models for direct marketing

Quantitative models employed for direct marketing purposes perform a series of activities essentially related to: (1) inputs, that is, the collection of consumer data; (2) processing, concerning the selection of target customers and prospects, as well as profiling, cross-selling, and upselling activities; and (iii) outputs, that is, the evaluation of results (Wasson, 2006).

As for consumer data, direct marketing models usually consider sociodemographics and purchase behaviour data, such as consumers' transactions, feedback, and web-browsing behaviour (Bose & Chen, 2009; Venugopal & Baets, 1994). Transaction data can be used for designing direct marketing promotion, through the classic recency, frequency, and monetary model (Malthouse, 1999; Piersma & Jonker, 2004; Suh, Noh, & Suh, 1999), which considers: the recency of the last purchase (the time elapsed since the last purchase, or the number of unfruitful solicitations); the purchase frequency (the number of purchases over time); and the monetary value of past responses (the money spent during a given period or for the last purchase). The accuracy of response models can be improved by processing additional data on consumers' purchase behaviour (Heilman, Kaefer, & Ramenofsky, 2003).

With respect to processing activities, several techniques are employed to identify target populations, such as statistical techniques and machine learning techniques. Traditionally, conventional compensatory linear models, based on basic and advanced statistical and econometric techniques, such as multiple linear regressions, discriminant analysis, and logit/probit are used to mimic individual evaluations, preferences, and choice processes (Levin & Zahavi, 1998; Malthouse, 1999; Van den Poel & Buckinx, 2005; West, Brockett, & Golden, 1997). The estimation of a fixed-form function of relationships between dependent and independent variables, as well as the assumptions on the distribution of input variables and error, are typical advantages of conventional statistical techniques; nevertheless, the non-linearity of preferences and the importance of non-compensatory rules for studying judgement and choice make them inadequate for investigating direct marketing issues. Furthermore, linear models can consider only a limited number of input variables and evaluate only a narrow range of different solutions. Therefore, the variety of constantly updated and frequently noisy customer databases does not allow researchers and practitioners to obtain valuable results.

In order to overcome these limitations, new methods and techniques have been recently proposed for designing market response models, which contain 'response/no-response' as the dichotomous dependent variable and consider as output variables the purchase occurring within a given time period, the purchase value, and the time elapsing between two purchases (Baesens, Viaene, Van den Poel, Vanthienen, & Dedene, 2002; Bhattacharyya, 1999; Kim & Street, 2004; Kim, Street, Russell, & Menczer, 2005; Levin & Zahavi, 1998). Specifically, machine learning techniques, such as ANNs and data-mining techniques, can be defined as 'processes in which computer programs can learn to improve their performance from experience of doing certain tasks' (Mitchell, 1997, p. 5). Initially built for science and medical applications, ANNs are now considered an alternative tool for studying direct mail marketing campaigns (Viaene et al., 2001; Zahavi & Levin, 1997; Zhang, Jansen, & Spink, 2009). Unlike statistical techniques, they merely require a flexible mathematical function defining the relationships between input and output variables; they can also be considered non-linear and non-parametric models, as their parameters are set up by employing a learning algorithm (White, 1992; Zahavi & Levin, 1997). These features allow them to overcome the limitations of linear multivariate statistical techniques, for which the functional form of the problem-solving equation must be known a priori and which are not capable of acquiring new information from firms' dynamic and competitive environments, with the consequent need for a periodical re-estimation of the models (Venugopal & Baets, 1994). Additionally, ANNs have a high modelling flexibility and adaptability, as they can control the learning process and adjust their parameters as new input data are available. Against these benefits, ANNs show a modest level of interpretability, as they can be used for the detection of significant independent variables but are not able to describe the relationships between inputs and outputs. Other limitations are related to the fact that they do not have a test statistic, are deficient of a theoretical background, and need time to process data (Vellido, Lisboa, & Vaughan, 1999). Furthermore, as ANNs can take many forms, the level of complexity can affect their stability; therefore, their architecture needs to be designed carefully (Judd, 1990). A description of ANNs is provided in the following section.

Artificial neural networks

ANNs are descriptive computational models designed for emulating the simplest functions and the properties of neurons of the human brain. Their aptitude to retrieve and learn complex samples of information from the problem data set, and the consequent ability to generalise that learned information, make them particularly appropriate for marketing research, especially in the area that formerly was the domain of multivariate statistical analysis (Bishop, 1995; Fausett, 1994; Haykin, 1999; Ripley, 1996).

Since the early 1980s, ANNs have increasingly been used in many disciplines, such as engineering, medical diagnosis, data mining, meteorology, and corporate business. Within business studies, ANNs have increasingly been employed in finance, management, and marketing. In finance, they have been instrumental in solving traditional issues such as bankruptcy prediction (Atiya, 2001; Muriel, 2006), financial market forecasting (De Leone, Marchitto, & Quaranta, 2006), and credit-scoring attribution (Quaranta, 2008, 2009), thus contributing to reducing losses and/or increasing profits of banks and financial companies. In business management, ANNs-aided analyses can improve the design of organisational structures, human resource management, and strategic decisions on new ventures (Quaddus & Khan, 2002).

In marketing, ANNs have typically been used for direct marketing purposes (Baesens et al., 2002; Cui & Wong, 2004; Curry, Davies, Evans, Moutinho, & Phillips, 2003; Ha, Cho, & MacLachlan, 2004; Kaefer et al., 2005; Kim et al., 2005; Venugopal & Baets, 1994; Zahavi & Levin, 1997), for studying customer satisfaction (Goode, Davies, Moutinho, & Jamal, 2005; Grønholdt & Martensen, 2005; Zhang et al., 2009), and predicting consumer shopping behaviour (Flynn, Eastman, & Newell, 1995; West et al., 1997) and promotion returns (Chee Wooi, & Kirikoshi, 2008; Lim & Kirikoshi, 2005). ANNs can also be employed for modelling individual responses to advertising messages (Curry & Moutinho, 1993) and brand choice (Hruschka, 2007), investigating the relationship between market orientation and performance (Silva, Moutinho, Coelho, & Marquez, 2009), estimating tourism demand (Kun-Huan, Moutinho, & Hui-Kuang, 2006), studying the effects of marketing planning on performances (Phillips, Davies, & Moutinho, 2001), developing new products (Thieme, Song, & Calantone, 2000), and accomplishing market basket analyses in retailing (Decker, 2005).

ANNs are self-learning and, generally, non-linear models trained to perform a function – an input/output map – from data. When a network is trained to execute a specific task, the system parameters are modified as circumstances change (training phase). Subsequently, the ANN parameters are set and the model is able to provide the output (testing phase). ANNs are therefore developed along a structured step-by-step process aimed at optimising a performance condition or satisfying particular, implicit internal restrictions (learning rule). A variety of ANNs having different characteristics and purposes is used for studying any, even complex, relationship between inputs (independent variables) and outputs (dependent variables) (Bishop, 1995; Fausett, 1994; Haykin, 1999; Ripley, 1996).

As for the training phase, although it is hard to make a clear-cut distinction, ANNs can be classified into two major types: supervised learning and unsupervised learning ANNs. Within the supervised learning type, both input variables and corresponding output values are identified and provided during the training, so that the ANN can fine-tune its parameters and try to match them with the desired response. Once a successful training is done, the ANN can be tested by presenting input data alone – without the desired results – in order to calculate an output value that approximates the target response (Bishop, 1995; Fausett, 1994; Rumelhart, Hinton, & Williams, 1986). Within the unsupervised learning type, the ANN contains some input data x and a function f, but in the training phase, output variables are not provided. The transfer function, which can correspond to any function of x and the network's desired output, is minimised on the basis of the specific application requirements (Hertz, Krogh, & Palmer, 1991; Kohonen, 1997). So, unsupervised learning is generally employed for estimation problems for which identified models are not available, whereas supervised learning is adopted for classification or regression tasks in which the desired output is expected.

The implementation of an ANN requires an optimal sample size, which has to be set up in relation not only to the size of the network, but also to the complexity of the function to be obtained and to the noise variance. This relates to the problem of the 'curse of dimensionality' – the exponential growth of hyper-volumes related to dimensionality (Bellman, 1961) – generally occurring in unsupervised learning. It means that the required sample has to increase non-linearly when the number of variables arises so that a small number of variables needs a large sample size. An undersized sample can be counterbalanced by creating an ensemble of networks, training them with a different re-sampling of the existing data, and then considering the average across the prediction of those networks. Furthermore, since a network with multiple output variables may face learning constraints, it can be useful to split it into a number of networks, one for each output, which are then combined into an ensemble. The architectural choices related to an ANN also involve the selection of useful and relevant input variables, a decision that can be made difficult by the curse of dimensionality, and the interdependency and redundancy of variables (Bishop, 1995).

One of the most common supervised learning ANNs is the Multilayer Perceptron (MLP), devised by Rumelhart and McClelland (1986) and made of a number of input-, hidden-, and output-layers, each containing a quantity of mutually interconnected neurons. Whereas the number of input and output nodes can be generally defined by the specific task, the number of hidden layers and units has to be carefully evaluated (Bishop, 1995; Haykin, 1999). One or multiple hidden layers are needed for linear and generalised linear models to process modest amounts of data with a great deal of noise, which limit the estimation of the non-linearities, while one hidden layer is sufficient for MLPs containing a number of continuous non-linear hidden-layer transfer functions (Bishop, 1995; Ripley, 1996). The number of hidden nodes should be set depending on the training sample size, the complexity of the activation function, and the amount of noise. If a form of regularisation is used, this step involves training various networks with different numbers of hidden nodes and selecting the solution with the minimum estimated generalised error. A supervised learning algorithm, in MLPs, allows the training phase to run, where data are used for adjusting the free parameters in order to minimise the prediction error, by comparing the resulting output with the desired outputs. One of the best known algorithms is the backpropagation algorithm (Fausett, 1994; Haykin, 1999; Patterson, 1996; Rumelhart et al., 1986; Werbos, 1994), which represents a gradient search procedure in which the gradient vector of the error surface is computed. The training phase has the purpose of locating the global minimum of this multidimensional error surface. The backpropagation algorithm is generally modified by the learning rate, usually chosen, according to the experiment, by the momentum term, and proceeds iteratively through a series of epochs. The training has the objective of generalising the results obtained, which is problematic in a situation of over-training, occurring when a network having too many weights shapes a complex function (Smith, 1996). A network with too many neurons in the hidden layer results in a scarce capability of generalisation, as it just learns one kind of examples with negative effects on model performances; conversely, a network with too few hidden nodes has a worse learning performance. This problem is better addressed by training different networks with different samples, and observing their individual performance. Besides a training and a validation sample, a test sample is also used in order to improve the confidence of the performance of the definitive model, and to make sure that the results obtained are valid (Moody, 1992; Smith, 1996).

Aims and objectives

The present study aims to verify the effectiveness of direct mail marketing campaigns by using ANNs and, specifically, a MLP ANN. This is achieved by pursuing the following objectives: (1) predicting the response rate of a direct mail marketing campaign and (2) identifying the characteristics and the purchasing intention of the campaign target.

A direct marketing campaign tries to maximise returns as positive responses out of a mailing list of customers and prospects (David Shepard Associates, 1999; Venugopal & Baets, 1994). The most important decisions regard targeting and prediction (Zahavi & Levin, 1997), as a successful direct marketing campaign is targeted to evaluate the attractiveness of active and non-active customers, to profile the best prospects, and to address to them the promotional elements that maximise returns (Kaefer et al., 2005; Venugopal & Baets, 1994).

Methodology

Both the mailing campaign and its target were simulated by using an experimental mailing campaign for book mail orders. A focus group involving eight students randomly selected in an Italian university was carried out in order to identify the type of product students would buy by mail order. Books were chosen because the purchase process does not entail direct observation of specific characteristics (binding, type of paper, text format). Subsequently, a pilot questionnaire was administered to 30 students in order to choose a genre (Guides/Handbooks) and then a specific book belonging to that genre (i.e. *Internet e il Web* by Rosario Viscardi).

The experimental mailing campaign was carried out by administering the main questionnaire to a random sample of 452 students. The questionnaire included the following items: (1) the last mail-order book purchase; (2) the first mail-order book purchase; (3) the frequency of mail-order book purchases; (4) the total amount of money spent on mail-order book purchases; (5) the number of books purchased for each different genre (essays, current events, crime novels, fiction, thriller, guides/handbooks); (6) sociodemographics (age and gender); and (7) the willingness to buy Rosario Viscardi's book by mail order. These variables were chosen as a result of previous studies found to be significant in explaining the response rate of direct marketing programs (Bose & Chen, 2009; Cui & Wong, 2004).

The software NeuroSolutions v.4.20 (NeuroDimension, Inc., Gainesville, FL) was used in order to build the MLP ANN model.

Data analysis and results

The implementation of the ANN model first requires the definition of input (independent variables) and output (dependent variables). The following input variables were chosen: (1) latest purchase; (2) first purchase; (3) purchase frequency; (4) total amount of money spent; (5) past purchases of essays; (6) past purchases of current event books; (7) past purchases of crime novels; (8) past purchases of fiction novels; (9) past purchases of thrillers; (10) past purchases of guides/handbooks; (11) age group 19–23 years; (12) age group >23 years; (13) male gender; and (14) female gender. The output variable was the 'willingness to buy/unwillingness to buy' the book *Internet e il Web*. The sample was split into four groups according to the following four possible combinations: mail order book 'purchase/non-purchase' and 'willingness to buy/unwillingness to buy' the book *Internet e il Web* by mail order. In order to obtain an ANN capable of providing an accurate forecasting, only 205 students included in the following groups were considered in the analysis: (1) students who purchased books by mail order and were willing to buy the proposed book and (2) students who purchased books by mail order and refused to buy the proposed book.

A cross-validation procedure was implemented in order to determine when the training process could stop, thus identifying the nature of the functional relation between input and output variables specified in the model. Data were partitioned in the following subsamples: (1) a training sample ($N = 124$; 60% of the total sample) used for the training of the ANN; (2) a validation sample ($N = 29$; 15% of the total sample) employed for the identification of the appropriate number of interactions, that is, how many times data should be considered in the ANN to update the connection weights; and (3) a testing sample or hold-out sample ($N = 50$, 25% of the total sample) used during the testing of the model when the ANN formulated the forecasting.

The ANN had the following structure: (1) an input layer containing 14 neurons; (2) a hidden layer containing four neurons; and (3) an output layer containing two neurons. The hyperbolic tangent function was chosen, and the parameters considered for the training process control were a momentum coefficient equal to .7, and a learning rate equal to 1 (both for the hidden and the output layer). The training of the model was accomplished by specifying the following parameters: the number of epochs equal to 1000, and the mean squared error (MSE) equal to .01. Furthermore, it was chosen to randomise the initial weights and to stop the cross-validation process after 100 epochs. The training process was run in order to minimise the MSE value. The second phase aimed at obtaining the best network: even if the MSE obtained for the training sample was smaller than the MSE obtained for the validation sample, the training process was stopped in the second phase in order to avoid the problem of overfitting, as shown in Table 1.

After the training indicated the best network, the testing (or validation) of the model was carried out to verify its prediction performance. The testing process was

Table 1 Report on the best training.

Best network	Training	Validation sample
Number of neurons in the hidden layer	10	8
Training cycle	1	3
Epochs	743	349
MSE minimum value	.060021177	.229726538
MSE final value	.069503948	.320999652

reiterated first on the testing sample, second on the training sample, and finally on the validation sample. The testing on the training sample obtained the best performance, as, in this sub-sample, the ANN scored the best connection weights, capable of minimising the MSE and producing an accurate prediction, as shown in Table 2. The ANN obtained an accurate prediction performance: 83% of correctly classified responses, an average of the three subsamples (76.5% for the testing sample, 92.5% for the training sample, and 79.5% for the validation sample), with a correlation coefficient equal to 62%, an average of the three subsamples.

A sensitivity test was implemented, in order to measure the importance of each input variable used in the model and believed to influence the intention to purchase the proposed book, by determining the weight of each variable on the corresponding output value. The sensitivity test returned 70% of correctly classified responses, an average of the three subsamples (63% for the testing sample, 73% for the training sample, and 75% for the validation sample), and showed that the following input variables influenced the purchase intention: past purchase of guide/handbooks (equal, on average, to .32); age group 19–23 years (equal, on average, to .17); male gender (equal, on average, to .16); female gender (equal, on average, to .11); and age group >23 years (equal, on average, to .07). The importance of the variables found to influence the purchase intention was estimated by a MRA and a LRA,in order to evaluate the predictive results of both.

As data considered in the ANN were analysed for each of the three subsamples, the MRA was estimated not only for the total sample, but also for the testing sample, the training sample, and the validation sample by considering the 'willingness to buy' as the dependent variable and the same independent variables considered in the ANN model. Results concerning the testing sample ($R = .65$; $R^2 = .42$; $F = 2.37$; $p > .05$), the training sample ($R = .33$; $R^2 = .11$; $F = 1.14$; $p > .05$), the validation sample ($R = .68$; $R^2 = .46$; $F = 1.30$; $p > .05$), and the total sample ($R = .35$; $R^2 = .12$;

Table 2 Report on the prediction performance.

Performance	Choice not to buy	Choice to buy
MSE (mean square error)	.206125972	.221618468
NMSE (normalised mean square error)	.868802874	.934102384
Min ABS Error (minimum absolute error)	.005238685	.004550099
Max ABS Error (maximum absolute error)	1.033821862	1.033197522
r (correlation coefficient)	.562960422	.54386825
Percent correct	68.42105103	91.66666412

$F = 2.32$; $p > .05$) showed that the independent variables did not statistically influence the 'willingness to buy'. Consequently, the ANN model had a higher predictive performance than the regression model.

Discussion

Results showed that the effectiveness of a direct marketing campaign can be improved by using an ANN able to predict the response rate and to outline the prospects' profile.

The MLP ANN was able to forecast the response rate, with 83% of exactly classified responses, on average, for the three subsamples. Testing the model allowed a comparison of the model's predicting performance against the real data in order to identify responses that the ANN had incorrectly classified as positive or negative about the choice expressed by individuals included in the sample as 'willing to buy'. The comparison showed that the prediction does not provide an explanation for the results obtained by the ANN. In fact, the ANN works as a 'black box' (Vellido et al., 1999) in that it only reproduces the behaviour of a system but does not explain its internal relations.

The sensitivity test, however, allowed the identification of the input variables influencing the interviewees' purchase intention. The similarity of the results obtained for the three subsamples justified a generalisation to the total target, and demonstrated the logic by which the ANN elaborated the data.

The input variable past purchase of guide/handbooks was the driver that most influenced the purchase intention of subjects included in the testing sample and the validation sample, whereas age group 19–23 years most influenced the individuals included in the training sample. Male gender affected the purchase intention of the total sample and particularly of the training sample. Female gender had a moderate influence on the purchase intention of the training sample. Another input variable, age group > 23, was also relevant to the choice of the proposed book for the three subsamples. The past purchase of a specific genre – such as thrillers, crime novels, and essays – had a marginal influence on the purchase intention. In spite of expectations, results showed a lack of significance of input variables concerning characteristics reasonably considered important for identifying prospects and usually investigated in the well-known RFM model, such as the latest purchase, the first purchase, the frequency of purchasing, and the total amount of money spent for book mail orders over the last year (Baesens et al., 2002; Kaefer et al., 2005; Kim & Street, 2004; Kim et al., 2005; Levin & Zahavi, 1998). Instead, a direct marketing mailing campaign for the promotion of a book such as *Internet e il Web*, or similar guides or handbooks, should be targeted to subjects having already purchased in the past the same genre. This result confirms the usefulness of including variables referring to product characteristics, as previous studies have highlighted (Bose & Chen, 2009). In addition, it will be appreciated more by males than females, and more by students within the age group 19–23 years than those older than 23 years. The scarce importance of input variables concerning past mail orders in the last year showed that the book could also be promoted to individuals regardless of their book purchases over the last 12 months. This can facilitate the promotion of this type of guide, as the direct marketing mail campaign can be efficiently addressed to male students aged 19–23 years.

Table 3 Fit of ANN, multiple regression analysis, and logistic regression analysis.

Testing sample			Validation sample			Training sample		
Best network	MRA	LRA	Best network	MRA	LRA	Best network	MRA	LRA
$r = 44\%$	$R^2 = 43\%$	$R = 41\%$	$r = 55\%$	$R^2 = 47\%$	$R = 52\%$	$r = 89\%$	$R^2 = 11\%$	$R^2 = 15\%$

Results obtained with the ANN model were then compared with MRA and LRA for the three considered subsamples, as shown in Table 3. Even if both MRA and LRA obtained a high performance for the testing sample and the validation sample, the high *p*-value ($p > .05$) excludes the significance of any variables included in the model and heavily limits the practical use of this methodology. Furthermore, results of LRA showed a very low rate of accurately classified responses: 56% for the training sample, 50% for the validation sample, and 36.7% for the testing sample.

The devised ANN model outperforms MRA and LRA and seems to be a robust and accurate tool, which is consistent with what previous studies have reported (Cui & Wong, 2004; Flynn et al., 1995; Goode et al., 2005; Grønholdt & Martensen, 2005; Lim & Kirikoshi, 2005; Thieme et al., 2000; West et al., 1997). This result demonstrates the improvements of the efficiency of a direct mail marketing campaign by splitting choices regarding willingness to buy according the two alternatives – positive or negative – and elaborating for each typology both the response rate and the weight of each input variable influencing the purchase intention.

Managerial implications

Prediction (i.e. anticipation of the response rate) and targeting are both key to decision making underlying direct marketing campaigns (Zahavi & Levin, 1995, p. 35). The results of this study demonstrate the usefulness ANNs for supporting direct marketing campaigns. The accuracy of prediction can help marketing practitioners deal with both prediction and targeting, as well as identify the variables that may affect response rate. This study offers a contribution for expanding the empirical horizons of neural methodologies; indeed, as Zhang, Patuwo, and Hu (1998, p. 56) state: 'although neural networks represent a promising alternative to traditional statistical methods, a long experimentation is required in order to confirm their ability to support predictions and to devise a modeling methodology which facilitates their setup and allows a more widespread use'. As shown in business and social science fields, such as sociology (Flynn et al., 1995) and consumer behaviour (Goode et al., 2005; Kim et al., 2005), this study confirms the results of recent research (Baesens et al., 2002; Bhattacharyya, 1999; Kaefer et al., 2005; Kim et al., 2005) and their implications for marketing managers. Direct marketing companies that exploit the advantages of information technologies formulate marketing strategies underpinned by the capability of predicting the purchase behaviour of customers and prospects belonging to relatively homogeneous segments. Consistent with this approach, the implementation of direct marketing initiatives could yield two orders of advantages: first, their ability to predict a specific target's purchase behaviour is fundamental for the formulation of marketing strategies; second, the spread of database marketing would even enhance the power and effectiveness of ANNs, as the huge quantity of data accumulated in firms' databases might be exploited to train models and support the development of data-intensive applications. At the same time, the use of new techniques for evaluating the effectiveness of direct mailing tackles the targeting problem as well, since the prediction output splits the sample population into two groups: the target, made up of prospect customers, and the non-target, made up of non-prospects. After identifying specific prospects, their willingness to purchase – calculated as a function of the relevant individual attributes (i.e. age, gender, location, occupation, etc.) – could be combined with their purchase history, thus generating valuable information for the design of future mailing campaigns. Due to these important advantages, the use of causal prediction models like ANNs show the capability of speeding up a transition from the traditional mass marketing paradigm, oriented to large market segments, to a one-to-one logic of relational marketing addressed to individual consumers. An example of this can be found within the context of e-commerce, where the practice of web customisation is crucial for marketers who try to deploy a CRM fully, while facing the challenges of the contemporary competitive environment (Allen, Kania, & Yaeckel, 2001). In this respect, ANN-aided direct mailing campaigns offer opportunities for even greater profitability, as they support a more effective and efficient web customisation and allow the identification of the most promising prospects for subsequent more successful direct mailing initiatives (Venugopal & Baets, 1994).

Conclusions

The design and the implementation of a direct marketing mailing campaign by using a MLP ANN allowed the general purpose of the study to be achieved – namely, to demonstrate the effectiveness of this tool for predicting response rate and profiling prospects and customers. To do so, a mere, although accurate, prediction of the individual response rate per se is not enough, but needs to be complemented by the exact indication of what determined that prediction. This leads to the most relevant outcome that the ANNs provided: a measure of the extent to which input variables – included in the model as possible determinants of the purchase intention – have influenced the actual purchase intention expressed by the target population through the questionnaire. The lack of a real direct marketing database has hindered the research process, since a data collection phase had to be experienced in order to track the purchase history of the target mail-order books. However, having operated on a real random target reinforces the validity of the research results and further emphasises the prediction accuracy of the ANN. Moreover, the a posteriori use of the neural tool for verifying the prospects' purchase intention is an additional application of the model, where the ANN's ability to detect the determinants of a given phenomenon is unchanged, if not even strengthened.

The very good prediction performance would encourage further applications of ANNs for investigating similar phenomena. For example, ANNs can be considered a valuable tool for studying classification, regression, or mapping tasks when, due to a considerable amount of input data, the research problem is not adequately known. ANNs are also advantageous in the presence of a non-linear or chaotic customer databases structure (Venugopal & Baets, 1994). Notwithstanding their simplicity, they should be meticulously designed, as their complexity impacts on performances. Indeed, an excessive complexity of the task or the data set characteristics can cause an overfitting of the model, which is sensitive to data noise. Unlike conventional statistical techniques, the use of ANNs can be problematic when a high level of interpretability is required. To solve these problems, several hybrid models have been proposed in the literature, which combine, for example, neural networks, logit models, and RFM models (Bose & Chen, 2009; Kim & Street, 2004; Suh et al., 1999; Zahavi & Levin, 1997).

Nevertheless, ANNs showed a capability of grounding on a narrow range of sample occurrences for predicting the future behaviour of a phenomenon and for exploiting the non-linearity within the sample in order to find out complex relationships among th edata. The MLP ANN designed for this study provided a prediction of the possible response rate for each of the target recipients, while the MRA demonstrated the superiority of the ANN as a non-parametric model for the identification of hidden relationships between input and output variables. Its ability to extrapolate implicit but potentially useful relationships among data within business databases characterises the ANN as a useful and appropriate tool for a deeper analysis of customer bases, aimed at predicting behaviours and anticipating trends.

References

Allen, C., Kania, D., & Yaeckel, B. (2001). *Internet world guide one to one marketing*. New York: John Wiley.

Atiya, A.F. (2001). Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results, Senior Member, IEEE.

Baesens, B., Viaene, S., Van den Poel, D., Vanthienen, J., & Dedene, G. (2002). Bayesian neural network learning for repeat purchase modelling in direct marketing. *European Journal of Operational Research*, *138*(1), 191–211.

Barwise, P., & Farley, J.U. (2005). The state of interactive marketing in seven countries: Interactive marketing come of age. *Journal of Interactive Marketing*, *19*(3), 67–80.

Bellman, R. (1961). *Adaptive control processes: A guided tour*. Princeton, NJ: Princeton University Press.

Bhattacharyya, S. (1999). Direct marketing performance using genetic algorithms. *Journal on Computing*, *11*(3), 248–257.

Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford University Press.

Bose, I., & Chen, X. (2009). Quantitative models for direct marketing: A review from a systems perspective. *European Journal of Operational Research*, *195*(1), 1–16.

Chee Wooi, L., & Kirikoshi, T. (2008). Understanding the effects of pharmaceutical promotion: A *neural network* approach guided by genetic algorithm–partial least square. *Health Care Management Science*, *11*(4), 359–372.

Cui, G., & Wong, M.L. (2004). Implementing neural network decision support in direct marketing. *International Journal of Market Research*, *46*(2), 235–254.

Curry, B., Davies, F., Evans, M., Moutinho, L., & Phillips, P. (2003). The Kohonen self-organising map as an alternative to cluster analysis: An application to direct marketing. *International Journal of Market Research*, *45*(2), 191–211.

Curry, B., & Moutinho, L. (1993). Neural networks in marketing: Modeling consumer responses to advertising stimuli. *European Journal of Marketing*, *27*(7), 5–20.

David Shepard Associates (1999). *The new direct marketing: How to implement a profit-driven database marketing strategy* (3rd ed.). Homewood, IL: Irwin.

Decker, R. (2005). Market basket analysis by means of a neural network. *International Review of Retail, Distribution & Consumer Research*, *15*(2), 151–169.

De Leone, R., Marchitto E., & Quaranta, A.G. (2006). Autoregression and artificial neural networks for financial market forecast. *Neural Network World*, *16*(2), 109–128.

Direct Marketing Association (2009, January). *Future of direct marketing*. DMA Report.

Drozdenko, R.G., & Drake, P.D. (2002). *Optimal database marketing*. Thousand Oaks, CA: Sage.

Fausett, L. (1994). *Fundamentals of neural networks*. New York: Prentice Hall.

Flynn, L.R., Eastman, J.L., & Newell, S.J. (1995). An exploratory study of the application of neural networks to marketing: Predicting rock music shopping behavior. *Journal of Marketing Theory and Practice*, *3*(2), 75–85.

Goode, M.M.H., Davies, F., Moutinho, L., & Jamal, A. (2005). Determining customer satisfaction from mobile phones: A neural network approach. *Journal of Marketing Management*, *21*(7), 755–778.

Grønholdt, L., & Martensen, A. (2005). Analyzing customer satisfaction data: A comparison of regression and artificial neural networks. *International Journal of Market Research*, *47*(2), 121–130.

Ha, K., Cho, S., & MacLachlan, D. (2004). Response models based on bagging neural network. *Journal of Interactive Marketing*, *19*(1), 17–30.

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Heilman, C.M., Kaefer, F., & Ramenofsky, S.D. (2003). Determining the appropriate amount of data for classifying consumers for direct marketing purposes. *Journal of Interactive Marketing*, *17*(3), 5–28.

Hertz, J., Krogh, A., & Palmer, R. (1991). *Introduction to the theory of neural computation*. Redwood City, CA: Addison Wesley.

Hruschka, H. (2007). Using a heterogeneous multinomial probit model with a neural net extension to model brand choice. *Journal of Forecasting*, *26*(2), 113–127.

Jacob Zahavi, J., & Levin, N. (1995). Issues and problems in applying neural computing to target marketing. *Journal of Direct Marketing*, *9*(3), 33–45.

Judd, J.S. (1990). *Neural network design and the complexity of learning*. Cambridge, MA: MIT Press.

Kaefer, F., Heilman, M.C., & Ramenofsky, D.S. (2005). A neural network application to consumer classification to improve the timing of direct marketing activities. *Computers and Operations Research*, *32*, 2595–2615.

Kim, Y.S., & Street, W.N. (2004). An intelligent system for customer targeting: A data mining approach. *Decision Support Systems*, *37*(2), 215–228.

Kim, Y.S., Street, W.N., Russell, G.J., & Menczer, F. (2005). Customer targeting: A neural network approach guided by genetic algorithms. *Management Science*, *51*(2), 264–276.

Kohonen, T. (1997). *Self-organizing maps*. Berlin: Springer-Verlag.

Kun-Huan, H., Moutinho, L., & Hui-Kuang, T. (2006). An advanced approach to forecasting tourism demand in Taiwan. *Journal of Travel and Tourism Management*, *21*(4), 15–24.

Levin, N., & Zahavi, J. (1998). Continuous predictive modeling – A comparative analysis. *Journal of Interactive Marketing*, *12*(2), 5–22.

Lim, C.W., & Kirikoshi, T. (2005). Predicting the effect of physician-directed promotion on prescription yield and sales uptake using neural networks. *Journal of Targeting, Measurement and Analysis for Marketing*, *13*(2), 156–167.

Malthouse, E.C. (1999). Ridge regression and direct marketing scoring models. *Journal of Interactive Marketing*, *13*(4), 10–23.

Mitchell, T. (1997). *Machine learning*. Boston: MCB/McGraw-Hill.

Moody, J.E. (1992). The effective number of parameters: An analysis of generalization and regularization in nonlinear learning systems. In J.E. Moody, S.J. Hanson, & R.P. Lippman (Eds.), *Advances in neural information processing systems 4*. San Mateo, CA: Morgan Kaufmann.

Muriel, P. (2006). Neural networks applications in bankruptcy forecasting: A state of the art. *Neural Computing and Applications*, *15*(2), 154–163.

Patterson, D. (1996). *Artificial neural networks*. Singapore: Prentice Hall.

Payne, A., & Frow, P. (2005). A strategic framework for customer relationship management. *Journal of Marketing*, *69*(4), 167–176.

Pepper, D., & Rogers, M. (1993). *The one to one future*. New York: Doubleday.

Phillips, P.A., Davies, F.M., & Moutinho, L. (2001). The interactive effects of strategic marketing planning and performance: A neural network analysis. *Journal of Marketing Management*, *17*(1/2), 159–182.

Piersma, N., & Jonker, J.J. (2004). Determining the optimal direct mailing frequency. *European Journal of Operational Research*, *158*(1), 173–182.

Quaddus, M.A., & Khan M.S. (2002). Evolution of artificial neural networks in business applications: An empirical investigation using a growth model. *International Journal of Management and Decision Making*, *3*(1), 19–34.

Quaranta, A.G. (2008). Attribuzione dello Scoring Aziendale nel Contesto Basilea 2. *Banche e Banchieri*, *2*, 125–137.

Quaranta, A.G. (2009). Modelli di Rating Interno e Propensione al Rischio del Management. *Banche e Banchieri*, *6*, 451–465.

Ripley, B.D. (1996). *Pattern recognition and neural networks*. Cambridge, England: Cambridge University Press.

Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning internal representation by error propagation. In D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel distributed processing explorations in the microstructure of cognition*, *Vol. 1: Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Rumelhart, D.E., & McClelland, J.L. oup. (1986). Parallel distributed processing. Explorations in the microstructure of cognition. Volume I. Cambridge, MA: MIT Press.

Silva, M., Moutinho, L. Coelho, A., & Marquez, A. (2009). Market orientation and performance: Modeling a neural network. *European Journal of Marketing*, *43*(3/4), 421–437.

Smith, M. (1996). *Neural networks for statistical modeling*. Boston: International Thomson Computer Press.

Suh, E., Noh, K.C., & Suh, C.K. (1999). Customer list segmentation using the combined response model. *Expert Systems with Applications*, *17*(2), 89–97.

Thieme, R.J., Song, M., & Calantone, R.J. (2000). Artificial neural network decision support systems for new product development project selection. *Journal of Marketing Research*, *37*, 499–507.

Van den Poel, D., & Buckinx, W. (2005). Predictive online-purchasing behavior. *European Journal of Operational Research*, *166*(2), 557–575.

Vellido, E., Lisboa, P.J.G., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). *Expert Systems with Applications*, *17*(1), 51–70.

Venugopal, V., & Baets, W. (1994). Neural networks and their applications in marketing management. *Journal of System Management*, *45*(9), 16–21.

Viaene, S., Baesens, B., Van Gestel, T., Suykens, J.A.K., Van den Poel, D., Vanthienen, J., De Moor, B., & Dedene, G. (2001). Knowledge discovery in a direct marketing case using least squares support vector machine. *International Journal of Intelligent System*, *16*(9), 1023–1036.

Wasson, C.S. (2006). *System analysis*, *design and development: Concept*, *principles and practices*. Hoboken, NJ: Wiley.

Werbos, P.J. (1994). *The roots of backpropagation*. New York: Wiley.

West, P.M., Brockett P.L., & Golden, L.L. (1997). A comparative analysis of neural networks and statistical methods for predicting consumer choice. *Marketing Science*, *16*(4), 370–391.

White, H. (1992). *Artificial neural networks: Approximation and learning theory*. Oxford, England: Blackwell.

Zahavi, J., & Levin, N., (1997). Applying neural computing to target marketing. *Journal of Direct Marketing*, *11*(1), 5–22.

Zhang, G., Jansen, B.J., & Spink, A. (2009). Identification of factors predicting clickthrough using neural network analysis. *Journal of the American Society for Information Science and Technology*, *60*(3), 557–570.

Zhang, G., Patuwo, B.E., & Hu, M.Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*(1), 35–62.

About the authors

Gianluigi Guido (PhD, University of Cambridge, UK) is full professor of marketing and place marketing at the Faculty of Economics of the University of Salento, Lecce, Italy, and professor of marketing research at the Faculty of Economics of the LUISS University of Rome. He has written 12 books and as many chapters in edited books. His publications in English include: *Behind Ethical Consumption* (Bern: Peter Lang, 2009); *Computer-Mediated Relationships and Trust* (eds. L.L. Brennan and V.E. Johnson), New York: Hershey, 2008; *The Salience of Marketing Stimuli*, Boston: Kluwer, 2001; *Readings in Global Marketing* (eds. M.R. Czinkota and I.A. Ronkainen, London: Dryden Press, 1995). He has published more than 120 articles in national and international journals, including *International Journal of Market Research*, *Journal of Business and Industrial Marketing, Journal of Consumer Marketing*, *Journal of Economic Psychology*, *Journal of Euromarketing*, *Journal of Marketing Communications*, *Journal of Product and Brand Management*, *Journal of Satisfaction, Dissatisfaction and Complaining Behavior*, *Long Range Planning*, *Recherche et Applications en Marketing*, and *Research Policy*.

Corresponding author : Prof. Gianluigi Guido, Facoltà di Economia, Università del Salento, Via per Monteroni, Palazzo Ecotekne, 73100 Lecce, Italy.

T +39 0832 332665
E gianluigi.guido@unisalento.it

M. Irene Prete holds a PhD in economics and quantitative methods for market analysis at the University of Salento, Lecce, Italy, and is research associate in marketing and place marketing at the same university. She also holds a Science Master (Master Recherche) in management (Science de Gestion), a specialization in marketing, at the Ecole Doctorale Economie, Gestione et Espace at the University of Paris XII, France. Her publications in English include *Computer-Mediated Relationships and Trust* (eds. L.L. Brennan and V.E. Johnson), New York: Hershey, 2008; and articles in international journals, in particular *Journal of Targeting, Measurement and Analysis for Marketing* and *Advances in Management*.

E mi.prete@economia.unile.it

Stefano Miraglia holds a degree in business and management with a major in marketing at the Faculty of Economics of the University of Salento, Lecce, Italy. He currently is a PhD candidate in business management at Imperial College Business School, London, UK, where he is also a research assistant within the ESRC/AIM funded research project 'Platforms for Innovation'. His research interests are the design and organisational implications of replication in product-service businesses, the provision of integrated solutions capital goods sectors, and the architecture of business and technological platforms. His current doctoral research concerns the replication of modular product-service systems and platform-like architectures in the IT industry.

E stefano.miraglia@imperial.ac.uk

Irma De Mare holds a laurea (graduate degree) in business and management, with a major in marketing, at the Faculty of Economics of the University of Salento, Lecce, Italy. Currently, she is a control manager in the marketing department of a multinational company division in Perugia, Italy.

E IrmaDeMare@leroymerlin.it