Association for Information Systems
AIS Electronic Library (AISeL)

MCIS 2009 Proceedings

Mediterranean Conference on Information Systems (MCIS)

2009

Building Classifiers For Detecting Products Missing From The Shelf

Dimitris Papakiriakopoulos
Athens University of Economics and Business, dpap@aueb.gr

BUILDING CLASSIFIERS FOR DETECTING PRODUCTS MISSING FROM THE SHELF

Papakiriakopoulos, Dimitris, Athens University of Economics and Business, 47 Evelpidon & Lefkados Str., 113 62 Athens, Greece, dpap@aueb.gr

Abstract

The problem of products missing from the shelf is a major one in the grocery retail sector, as it leads to lost sales and decreased consumer loyalty. Yet, the possibilities for detecting and measuring an out of shelf situation are limited, mainly conducted via a visual shelf check. The existence of a method for detecting the products that are not on the shelf based on sales data would be valuable, offering an accurate view of the shelf availability both to retailer and the product suppliers. In this paper, we suggest a method based on the employment of machine learning techniques, in order to develop a rule based system. Results up to now presents that rules related with the detection of out of the shelf products are characterized by acceptable levels of accuracy.

Keywords: *Out Of Stock, Classification Problem, Supply Chain Management, Retailing*

1 INTRODUCTION

Consumer value and satisfaction are fundamental to building consumer loyalty (to the brand) and shopper loyalty (to the store) and to increase sales and category profitability (Colacchio et al. 2003). A powerful way to create value and satisfaction is to keep shelves fully ranged (Roland Berger 2002), but out of shelf (OOS) is still a frequent phenomenon in the grocery retail sector. Out of shelf rates vary wildly among retailers and their outlets depending on a variety of factors, but the majority tends to fall in the range of 5-10 percent. In their analysis, which is a compilation of many global surveys on the extent, causes, and consumer responses to retail out of shelf situations in the grocery retail sector, Gruen et al. (2002) estimate an overall average OOS rate of 8.3 percent.

However, in most European countries levels between 10 and 15 percent are not unusual (Roland Berger 2003). Emmelhainz et al.'s (1991) research results show, for instance, that a stock-out can make a manufacturer lose more than half of his buyers to competitors, whereas retailers face the loss of up to 14% of the buyers of the missing product. This revenue loss (approximates 1.5% of sales) not only stems from lost product sales during the OOS period, but can also extend to later periods or other product categories (Campo et al. 2000).

In this paper, we investigate the possibility of developing a method that detects the OOS products, utilizing discrimination / classification techniques. In more detail, having available the sales data, ordering info, product assortment of the store etc. we study the development of a rule based system that will automatically discover OOS situations on a daily basis for all the stores of a retail chain. The next section briefly presents the related literature regarding the OOS problem and the research methodology follows. The next section provides details regarding the development of the rules which are the classifier instrument of the proposed detection system. The paper continues with a section referring to the actions undertaken for validation purposes. The last section shortly examines the findings and prescribes future issues.

2 RELATED WORK

2.1 Causes of the Out of Shelf problem

The term *"out-of-shelf"* (OOS) is used in grocery retailing to describe the situation where a consumer does not find the product he/she wishes to purchase on the shelf of a supermarket during a shopping trip.

Despite the extended literature on consumer reactions to out of shelf situations, very little has been written on the reasons behind the problem. In the relevant texts that are available (which are fairly sparse and largely empirical), we see a classification of the causes of OOS into two major areas (Gruen et al., 2002; Vuyk, 2003):

- *Retail store replenishment causes*, i.e. the product was not ordered or the ordered quantity was not enough to meet the actual consumer demand. Apart from the ordering parameters this category also implies and the shelving replenishment practices utilized by the store. (e.g. shelf-space allocation, shelf-replenishment frequencies, store personnel capacity etc.)
- *Combined upstream causes*, referring to the product was not delivered due to out of stock situations or other problems with the retailer's distribution centre (for centralized deliveries) or the supplier (for direct-store-deliveries). Other upstream causes are the delivery of the wrong product, and the delivery of smaller quantity of products.

The Out Of Shelf problem is related with stock-out, where the later is used in the pertinent literature to describe both the situations where the product does not exist in the store. In general a Stock Out certainty implies an OOS situation, while the opposite is not always stands. On the one hand the Stock Out problem has been investigated in the area of Inventory Management for over thirty years and several models has been presented. On the other hand, the OOS problem is mainly discussed in the marketing literature from the consumer reaction perspective (Campo et al. 2000).

Combining prior knowledge of studying Stock Out with the essential characteristics of the OOS problem, Table 1 summarizes the variables affecting the product availability in the store. The first column is the name of the variable which is further modelled as a specific attribute (e.g. sales velocity could be calculated as mean sales of the product for a period, but it can also be expressed in terms of how frequently a product is sold). The second column depicts the relevant work and the last column shows the relation between the variable and the related problem.

Variable	**Reference**	**Problem addressing**
Sales Velocity	(Anupindi et.al;1998)	Stock Out / Out of shelf
Inventory Level	(Clark and Lee 2000), (Downs et.al 2001)	Stock Out
Promotional product	(Gruen et.al 2002)	Stock Out / Out of shelf
Shelf space	(Yang 2001) (Desmet et.al 1998) (Corstjens and Corstjens 1999) (Urban, 1998)	Out of shelf
Stock Centralization	(Cetinkaya et al, 2000) (Nahmias and Smith 1994)	Stock Out
Market share	(Bell and Fiztsimons 2000)	Stock Out / Out of shelf
Seasonality	(Metters 1998)	Stock Out
Day	(Gruen 2002)	Out of shelf
Store size	(Gruen 2002)	Stock Out / Out of shelf
Employees	(Vyuk,2003)	Out of shelf
Store managers decisions	(Campo, 2004)	Stock Out Out of shelf

Table 21. Variables related to the problem

However we argue that through the intervention of AI techniques it is possible to develop an adaptive system for detecting the products that are not on the shelf. In more detail this research objective is to develop a system in the retailer side where at a daily base will check product sales and ordering information for every store and produce a list with the products that are not on the shelf. This list would be delivered to (a) the store manager in order to take corrective actions (b) to the account manager of the retail chain in order to have an accurate view regarding the product availability performance for every store and (c) to the product suppliers in order to inform them regarding the status of their products at every store.

In order to develop the required application, we initiate our effort by proposing a research model that recognizes the independent variables (also referred as attributes) and use it as a functional tool to the Knowledge Discovery in Database process (KDD) (Fayyad et al, 1996). The execution of the KDD process introduces several issues regarding the available data, the selection of the appropriate classification scheme etc. The result of the KDD is a set of rules to be applied and validated. All the aforementioned issues are discussed in the next sections.

3 RESEARCH METHODOLOGY

For the purpose of the study, the OOS has been formulated as a classification problem, where the class variable (Shelf Availability) has two mutually exclusive states. The first state describes the situation where the product exists on the shelf (EXIST), while the opposite is the OOS, indicating that the product is not on the shelf. The objective is to discriminate the EXIST from the OOS cases, through the utilization of appropriate classification algorithms (e.g. Neural Networks, Decision Trees etc). A common approach to handle classification problems is the Knowledge Discovery Process. The next summarizes the main actions undertaken within the scope of the KDD process

Data Selection: We selected nine representative stores from the same retail chain, based on their size (Small, Medium and Large) and conduct a physical OOS survey. The sample product list included 110 items selected with the method of stratified clustering. Sample products were from several categories like Shampoo, Diapers, Coffee, and Laundry etc. Through store visits a product availability list has been produced. The data gathered in order to support the research model are the Point-Of -Sales (POS) data from the nine stores and for all the products, the product assortment of every store for the week of the study and category of the products as maintained by the retailer as well as the ordering history of all the products for the nine sample stores.

Data cleaning and preprocessing: In order to support the research model various calculations had to be done. However the fluctuation of the retail operations made the data cleaning procedures essential for the progress of the research. For instance due to limited shelf space, some products were partially removed from the stores during Christmas holidays and were replaced by highly seasonal items. However this reduction was not reported, since it is a common practice. Thus with the data cleaning procedures we identified sales lags and take them into account while calculating the variables of the problem. The result was to calculate more accurately variables related with the sales velocity and the inventory levels. The list of the calculated independent variables are depicting in Table 2.

Variable	Description	Type
Sales Velocity	Sales average and variation	Num
	Daily sales average and variation	Num
	Frequency daily sales and variation	Num
	Average and variance zeroes days intervals between	Num

	sales	
	Index of sales velocity	Num
	Days from the last sale	Num
	Number of units sold today	Num
Inventory Level	Estimate the level of the unsold items and decide for high or low inventory	Nom
	Days from the last order	Num
Promotional product	Find products with discounts or products with in store promotion	Nom
Shelf space	Calculate the shelf layout.	Num
Stock Centralization	Explore if a product is usually ordered though the central warehouse or delivered directly by the supplier	Nom
Market share	Examine the importance of a product within its category.	Num
Seasonality	Utilizing Winters method decide whether or not the category is subject to seasonal effects.	Nom
Day	which is the current day	Nom
Store size	Depending on sales and product assortment decide whether the store is small, medium or large	Nom
Employees	The number of staff of the store at a daily base	Num
Store managers decisions	Decide when and how much to order Decide the product assortment	Nom

Table 22. List of independent variables

The thoroughly examination of the initial data set show the following characteristics

- Combination of numerical and nominal independent variables and a nominal class variable. Moreover missing values for some variables set, because of imperfect information.
- Noisy data due to the dynamic nature of retail business, caused by the frequent changes in the product assortments at each store, the seasonality of some products (even during past periods) and the existence of in store promotions and advertisement products etc.
- Imbalance class problem because the products exists on the shelf were more than the products missing.

Based on these it was realized to develop more training sets in order to tackle the aforementioned problems. As a result four different training sets had been developed. This are

- **Training Set 1** (TS1): It is the original training set and it is described by high noise on the data. Although it is not ideal for building classifiers, it provides an overall view regarding the stability and the response of a classification method to the OOS problem. Having in mind that in real life the system would work with such noisy data, we decided to keep this training set for monitoring purposes.
- **Training Set 2** (TS2): The second set of data was derived from the above by removing de-listed products, which are the items that seems to be in the product mix, but they had never sold from the store for the last six months. The OOS rate of the set was close to 5%. Although the set is much closer to the expected OOS rate for the specific retail chain, it still suffers from the imbalance class problem (Kubat and Matwin 1997)

- **Training Set 3** (TS3): The next set was based on the TS2 and the application of resampling technique had been conducted, biased to the OOS class (8%) and at the same time we increased the size of the set about 50%. The idea is to raise the OOS class closer to the average worldwide out of shelf rate in order to tackle the imbalance problem (Japkowicz 2000). Trough data resampling it is possible to measure the maximum theoretical classification effectiveness for every method, since balanced data are usually a prerequisite to apply the methods. However using this research strategy, would make the validation of the system a problem, since the ensemble classification model developed by such balanced data would fail when applied to the new imbalanced during the system testing.
- **Training Set 4** (TS4): This set come up from the TS2. In more detail we thoroughly examine the products that were reported as OOS for over 30 days, and at the same time after 3 days they were available. We assume that these phenomena deal the Hawthrone effect and downsizing the OOS problem. We consider that TS4 is depicts efficiently the real OOS situations.

Having a few training sets with different characteristics we moved to the next of the KDD process (Fayyad et al, 1996).

Data transformation: The transformation of the data had been in line with the research model. The available data showed that the independent variables like *Shelf space*, *Employees* and *Store's Managers Decisions* could not be supported from the available data provided by the retail chain, thus they were not utilized and removed. For example we didn't receive any data regarding how many employees were in the store, the role of every employee, the working hours etc.

Data mining: The selection of the variables related with the problem and the appropriate classification algorithm had been the major issues of this task. The selection of the variables had been based on the RelieF attribute ranking method (Kira and Rendell, 1992). After reducing the variables of the problem for every training set, we had to select the best classification scheme. In doing so we selected 14 different algorithms for classification and categorized to Statistical (e.g. Naïve Bayes, Logistic Discriminant) , Decision Tree (e.g. C4.5, Alternating Decision Tree) and Rule-Based algorithms (e.g. RIPPER, RIDOR). For the Decision Trees pre-pruning techniques were used in order to examine the resistance to the noisy data. The comparison of the classification algorithms had been done using the 10x2 Cross-Validated Paired t-test (Dietterich, 1996). The result of this step is on the one hand the variables of the problem and on the other hand the classification algorithm that best fits the data. The application of the algorithm is a set of rules that predicts the OOS situations.

Interpretation / Evaluation: The final step of the data mining process is divided into two parts. The first part is the selection / discussion of the rules by experts, while the second part includes the application of the rules to the real business practice, thus discussing the external validity of the selected algorithms.

4 DEVELOPING THE CLASSIFIERS

4.1 Classification algorithm selection

In the literature few classification algorithms exists and their performance depends on the nature of the problem. The employment of the 10x2 Cross-Validated Paired t-test has been the comparison method. The selected significance level has been very small (a=.001) as suggested by the literature (Salzberg, 1997). Usually the main attribute to decide if one algorithm performs better than the other is judged with the measure of accuracy. In our case we examined the accuracy abilities of the algorithms and found that the average accuracy level is about 88%, which could be described as high. However the high level of accuracy was not caused due to the good fit between the data and the classification method, but it occurred due to the fact that one class (EXISTS) of the problem was dominating. Moreover the accuracy levels between the four training sets were significantly different. Table 3 presents the three best algorithms for the correspondent training set.

TS1	TS2	TS3	TS4
RIPPER (91,3%)	Decision Table (87,4%)	C4.5 (93,7%)	C4.5 (83,5%)
C.4.5 (90,9%)	Random Forest (86,3%)	Decision Table (92,8%)	RIPPER (82,1%)
Alt.Dec.Tree (90,4%)	RIPPER (85,6%)	Random Forest (92,6%)	Decision Table (81,6%)

Table 23. Classification accuracy for different TS

Most of the best performing classification algorithms are structuring decision trees with variable accuracy across different training sets, while the statistical classification algorithms do not perform relatively well. Since the TS4 is considered to be closer to the real world, it sets the theoretical upper bound regarding how many OOS situations could be discovered through the utilization of a rule-based system. In practise, the accuracy of discovering OOS situations is calculated around 55%. This means only half of the OOS occurrences would be detected, which is considered as important, because it could increase the sales of the retail chain by 0,7% in the short run.

Apart from the accuracy, the reliability performance of the classification process is an important success factor. We consider an algorithm reliable when it does not predict an EXISTS occurrence as OOS, which implies low False Positive (FP) rate. During the experiments we observed that most of the algorithms had very good reliability performance. On the one hand, focusing on the minimum FP rate is a good option for validation purposes, because the classifiers would not tend to characterize existing products as OOS. On the other hand the system would lose the some of detection capabilities. This implies that during the development phase it was selected to have a bias system, while the other option is to increase the variance, because it is easier to validate a list with 80-100 products daily at every store, than having an extended (more than 250 products) list. Table 4 illustrates the best three classification algorithms for every training set. The comparison was based on t-paired test and the selected measure was the FP rate, which is included in the parenthesis.

TS1	TS2	TS3	TS4
Decision Table (0,046)	Logistic Model Tree (0,07)	C.4.5 (0.044)	C.4.5 (0,025)
Ripple Down (0,051)	Naïve Bayes Tree (0,085)	Decision Table (0.052)	Random Forest (0,027)
RIPPER (0,063)	Random Forest (0,091)	Naïve Bayes Tree (0.064)	RIPPER (0,03)

Table 24. Classification FP rate for different TS

In general most of the classification algorithms utilized in the experiments were not tending to classify EXISTS cases as OOS and promising reliable results. Note that from TS1 to TS4 the FP rate is getting smaller, except the TS2 where there are very few instances of OOS. In practise reliability level should be higher than

85%, which is acceptable by the users. To this end the classification algorithms seems to have a good accuracy and high reliability.

The study of accuracy and realibility of the classification algorithms is a mandatory step in order to ensure that these methods are applicable in the problem of the OOS, but the remaining question is which algorithms are best for every training set. In doing so, the classification algorithms examined with 10x2 Cross-Validated Paired using the F-Measure (Van Rijsbergen, 1979). The selection of the F-Measure based on the idea that it is a good trade-off between accuracy and reliability. The comparison between the algorithms based on the hypothesis that the "*x* algorithm over performs the *z* algorithm using the F-measure" as a comparison criteria at significance level a=.01. The result of such repetitive process forms Table 5, where each algorithm described with wins, when the algorithm is better than another, draws, when the algorithm has no statistical significance difference with another, and loses in the cases that the algorithm is worse.

Classification Algorithm \ **Training Set**	**TS1**	**TS2**	**TS3**	**TS4**
Statistical and Mathematical Algorithms				
Bayes Networks	(2,13,0)	(3,8,4)	(5,1,9)	(3,6,6)
Naïve Bayes	(0,0,15)	(2,3,10)	(1,3,11)	(2,2,11)
Logit	(3,12,0)	(2,8,5)	(1,3,11)	(3,5,7)
Support Vectors	(1,14,0)	(0,11,4)	(1,3,11)	(2,4,9)
Instance Based Algorithms				
Instance Base-k	(1,14,8)	(0,2,13)	(0,1,14)	(1,0,14)
K*	(3,12,0)	(8,7,0)	(9,6,0)	(8,7,0)
Decision Trees				
AD Tree	(1,14,0)	(3,10,2)	(5,1,9)	(4,9,2)
C4.5	(2,13,0)	(7,8,0)	(8,6,1)	(6,9,0)
Logistic Model Tree	(2,13,0)	(3,12,0)	(8,7,0)	(6,9,0)
NB Tree	(2,13,0)	(4,11,0)	(8,7,0)	(6,9,0)
Random Forest	(2,13,0)	(8,7,0)	(11,4,0)	(8,7,0)
Rule Based				
Decision Table	(2,13,0)	(5,10,0)	(8,6,1)	(4,11,0)
RIPPER	(1,14,0)	(4,11,0)	(8,5,2)	(6,9,0)
RIDOR	(1,14,0)	(2,13,0)	(7,8,0)	(2,13,0)
Neural Networks				
MLP	(1,14,0)	(3,9,3)	(7,1,7)	(5,8,2)
Radial Basis Functions	(1,14,2)	(0,2,13)	(0,4,11)	(0,0,15)

Table 25. Comparing classification algorithms with the F-Measure

The most stable categories of algorithms regarding the F-Measure are Decision Trees and Rule Based. Statistical classification algorithms are not appropriate for the OOS problem and this occurred due to the imbalance of the data sets. Although Instance based algorithms have the same construction mechanism for classifier, the comparison results show that K* is far better than Instance Base-k algorithm. However K* performs good enough because of noise removal. However in real world problems, removing the noise is a very complex task, thus the results obtained for the two Instance based algorithms found to be problematic and not promising. Finally from Neural Networks, only Multi Layer Perceptron has a good response to the problem. However the main criticism of Neural network topologies is the operation they provide as "black boxes", thus they would have limited contribution in the better understanding of the problem. Thus utilizing Decision Trees and Rule based algorithms is a reasonable research approach for the OOS problem.

4.2 Rules Selection

The identification of the accuracy and the reliability of the algorithms acted as a compass in the selection process of the rules. Instead of selecting a single decision tree, the approach to build an ensemble was selected in order to increase the detection capabilities of the system. Having the most accurate and reliable algorithms for every training set we had to make a "fair" mix of rules. Initially we got over 400 rules referring only for the OOS detection, which had been considered large for the validation purposes, although in some cases raises the question of overfitting the data. In doing a selection algorithm was designed in order to extract single rules from the decision trees. The algorithm is listed in the next

```
Building the OOS ensemble algorithm

Input: DecisionTree i, Threshold  T

For k=1 to 10
 Create Test Data Set k
    For i=1 to T
       R i ← Select the Braches of the Tree Labeled as OOS
       Accuracy(R i) ← True OOS/(True OOS + False OOS)

 Expected Accuracy(R i) ← Calculate Average Accuracy(R i)
 IF Expected Accuracy(R i) > T
    R ← R + R i

Output: Rule Set R
```

Figure 75. The ensemble algorithm for selecting the rules

The algorithm had as input the Decision Trees and the desired accuracy threshold T. The creation of random Test Data Sets, derived from the TS2 and TS4 because they found to be as the more realistic. The algorithm was initiated 10 times in order to examine how every single rule behaves when the test data set changes. By having the average accuracy of every single rule, it was possible to compare the rules and rank them based on the expected accuracy. At last 127 rules were selected having expected accuracy greater then 80%. The result of the selection algorithms populates the *R* set of rules, which are stored in the knowledge base of the system.

5 SYSTEM VALIDITY

The validation of the system was based on physical counts. The initial objective was to use the OOS lists (a list of products that according to the system are not available in the store), in order to calculate the support and confidence. However this task could not be accomplished due to the fact that information sources of the retailer were problematic. In more detail each store maintains a list of products (called Product Assortment or Product Mix) which supported by the store. However this information source is inaccurate because it includes few products that had been in the store and currently are not supported (e.g. promotional products, seasonal items etc). The variation between the number of records in the product assortment and the real capacity of the store (means how many different items are in the store) is very high and for small stores it could be 2.000 product codes, while in large stores the 10.000 product code difference is not an exception.

Thus we used an alternative path and at the first stage we made few physical audits in 6 different retail outlets. The procedure was the following

1. Visit the store
2. Use random walkthroughs and discover OOS products
3. Write the codes and inspect in detail the whole category.

After 15 days we collected a list of OOS single cases (more than 2000 counts) and expand the selection with EXISTS cases as derived from the POS Data, forming a single test set (TeS1 – 28.500 counts). . Note that for every one OOS case we were adding 10-14 different exist cases, in order to maintain the distribution close to 8% which is the world average OOS rate. Based on the idea that if a product had been mentioned OOS for a certain day, then this product would have been OOS for all the days before. Thus we expanded the OOS cases and following the procedure of inserting the EXISTS cases a larger test set (TeS2- 60.000 counts) derived. To this end we activated the system to produce the OOS lists for the 15 days of the trial and compare the detection results of the system with the test sets, allowing the computation of *confidence* and *support* measures for every rule.

Confidence is a proportion of how many cases detected right from the total number of cases detected, and profoundly is an indicator of the system's accuracy. *Support* measure describes the number of OOS cases detected right divided by the total number of available OOS cases. This measure describes the coverage of the solution. For the different test sets (TeS1 and TeS2) we found that 46 and 35 rules respectively where extremely accurate (Confidence = 100%) and covering 26% and 29% of the OOS cases. By lowering the confidence level it is expected more rules to participate and increase the support. Table 6 summarizes the findings for the 2 different (and related) test sets.

Test Set	Confidence Level	# of rules	Confidence	Support
TeS1	=1	46	1	0,268
	[0.8 , 1]	50	0,95	0,27
	[0.5 , 1]	56	0,77	0,279
TeS2	=1	35	1	0,297
	[0.8 , 1]	43	0,91	0,315
	[0.5 , 1]	56	0,742	0,34

Table 26. The confidence and support measures for the overall system.

Based on the table we can keep only 50 rules and at a 95% confidence level the system could find almost 27% of the total OOS cases occurring daily in the store. This result is likely high considering the diversity of OOS cases and the complexity of the problem. However the most interesting part is that the rules had been developed in different training set, were able to detect the OOS cases in a new and totally unknown test set, and seems that through the adoption of a rule based system the detection of OOS cases is possible.

Some rule examples depict in the next table. The first rule (Rule21) characterizes as OOS products that didn't sale for the last three days (LastPosDays >=3), the date of detection is Wednesday (day = 'Wednesday'), the area of interest is only the large stores of the retail chain (Store_Size = 'Large'), the standard deviation of sales only for Wednesday should be low (SD_DailyPosAvg <= 2.82) and finally the products are close to fast moving sales item (FastMovingIdx > 0.76). This rule has a relative high confidence but refers only to a small proportion of the total OOS occurring. It is high complex and very difficult for an expert of the industry to interpret. However having a closer look to Rule21 it is possible to argue that it detects products haven't make any sales from Saturday in large stores. In more detail having a product with small standard deviation on sales (controlled by the precondition of the SD_DailyPosAvg attribute) and this is almost a high frequent selling item (controlled by the FastMovingIdx attribute), it is rational to argue that this rule in order to achieve high confidence it prefers to wait for three days (so it is Wednesday) in order to characterize a product as OOS. Similar conclusions might be drawn from Rule43 and Rule47.

RuleID	**Rule Body**	**Confidence**	**Support**
Rule21	(LastPosDays >= 3) AND (day = 'Wednesday') AND (Store_Size = 'Large') AND (SD_DailyPosAvg <= 2.82) AND (FastMovingIdx > 0.76)	0,82	0,004
Rule43	(LastPosDays > 6) AND (SD_PosAvg > 7.9) AND (day = 'Tuesday')	0,42	0,01
Rule47	(TypeOfProducts = 'ADV')AND (Last_Order > 12) AND AND (Mean_Order_quantity < 6) AND (posavg> 1.9)	0,91	0,001

Table 7. Indicative rules examined.

From the empirical work the next was found

- Experts are able to make similar rules with low complexity form (such as rule43) but the confidence level is expected to be low.
- The nature of the OOS problem itself, different location of the stores, frequent changes in product assortments, promotional and seasonality effects, are important obstacles for having a small compact set of rules to cover all the OOS instances.

The suggested system was also compared with the EOI (European Out Of Shelf Index) for benchmarking reasons. After a joint effort of retailers and suppliers in the European grocery retail sector, and is referred to as the OOS Index. Taking into account only fast moving items with low sales volatility, the OOS Index monitors the sales of the corresponding products on a daily basis; if for a given day a product sells zero items (or lower than a predefined ceiling) then it is considered to be OOS. Using the TeS1 and TeS2 we found that the EOI has 36% accuracy and 0.27% support. In fact EOI is a good indicator only for days with high sales volumes (e.g Saturday), but the coverage of the solution provides (as expressed by the support measure) is really low. To this end the suggested system over-performs the existing European standard.

6 DISCUSSIONS AND FURTHER WORK

The initial result of this research shows that the application of AI techniques is able to increase the profitability of the retail chain, by downsizing the rate of OOS products. However we consider that the initial prototype needs to be further improved. Most of the accurate and reliable rules are based on the fact that if a product is not selling for a long period, then it is predicted as OOS. For example if a product with high velocity won't sale for 4 days, the fifth day would be detected as OOS so a next step for the system is to minimize the latency, through the incorporation of more reliable variables regarding the inventory levels, better handling of the data noise etc.

At the moment the detection system was tested with four different retail chains following the same method of work. Every retail chain acquired a different system set up based on the available data. The observed results regarding support and confidence were close to the aforementioned. From technical perspective few open issues are identified, like the development of score functions for the rules, the development of sophisticated attributes, the creation of clusters between different stores etc.

Launching the system in production raises issues related with the system is the user's acceptance and adoption. The stores managers seem to not feel comfortable with such a monitor system, because it could identify inefficiencies of the store's operations. However the early detection of the missing products could assists their duties and reduces the physical inspection of the shelf. Moreover it is expected that the system would affect the Out of shelf cases, thus new rules required and the design of the maintenance phase need also to taken into account.

References

Anupindi, R., Dada,M. and Gupta,S. (1998) Estimation of Consumer Demand with Stock out based Substitution: An Application to Vending Machine Products, Marketing Science, Vol.17, 406-23.

Bell, D. and Fitzsimons, G. (2000) An Experimental and Empirical Analysis of Consumer Response to Stockouts, Working Paper #00-001, Wharton Marketing Working Papers Series The Wharton School, University of Pennsylvania, Philadelphia, PA.

Campo,K. Gijsbrechts,E. and Nisol,P. (2000) Towards understanding consumer response to stock-outs. Journal of Retail, Vol76, No.2,pp.219-242.

Campo,K. Gijbrechts,E. και Nisol,P. (2004) Dynamics in consumer response to product unavailability: do stock-out reactions signal response to permanent assortment reductions?, Journal of Business Research, 57, pp.834-843

Cetinkaya, S., Lee, C.Y. (2000) Stock replenishment and shipment scheduling for vendor managed inventory. Management Science, Vol.46, No.2, pp.217-232.

Clark, T., Lee, H. (2000) Performance interdependence and coordination in business-to-business electronic commerce and supply chain management. Information Technology and Management, 1, 1/2, 85-105.

Colacchio, F., Tikhonova, O., Kisis, J. (2003) Consumer Response to Out-Of-Stock:Decision-making process and influencing factors. ECR Europe Berlin Conference

Corstjens, J., and Corstjens,M. (1999) Store wars: the battle for mindspace and shelfspace, Wiley, New York

Desmet, P., Renaudin.V, (1998). Estimation of product category sales responsiveness to allocated shelf space. International Journal of. Marketing Research, vol. 15, pp 443-457.

Dietterich, T. (1998) Approximate Statistical Tests for Comparing Supervised Classifications Learning Algorithms, Neural Computation 10, pp 1895-1923

Downs, B., Metters, R., Semple, J., (2001) Managing inventory with multiple products, lags in delivery, resource constraints, and lost sales: A mathematical programming approach, Management Science, Vol. 47, No. 3, pp.464–479.

Emmelhainz, M., Emmelhainz, L., Stock, J. (1991) Consumer Responses to Retail Stock-outs, Journal of Retailing, Vol. 67, No.2, pp.138–147.

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., (1996) From Data Mining to knowledge discovery in databases, Artificial Intelligence Magazine, Fall 1996, pp.37-54

Gruen, T., Corsten, D., Bharadwaj, S. (2002) Retail Out-of-Stocks: A Worldwide examination of Extent Causes and Consumer Responses. The Food Institute Forum (CIES, FMI, GMA)

Japkowicz,N., (2000) Learning from imbalance data sets, Proceedings on AAAI' 2000 Workshop on Learning from imbalance data sets

Kira,K. and Rendell,L. (1992) A practical approach to feature selection, Proceedings of the Ninth International Conference on Machine Learning, pp. 249-256, Morgan Kaufmann.

Kubat,M., Matwin,S., (1997) Addressing the curse of imbalance data sets: One sided sampling, Proceedings of the 14th International conference on Machine Learning, pp.179-186

Metters,R., (1998) Producing multiple products with stochastic seasonal demand and capacity limits, Journal of Operational Research Society, Vol. 49, No. 3, pp. 263-272

Nahmias,S. and Smith, S. (1994) Optimizing inventory levels in a two-echelon retailer system with partial lost sales, Management Science, Vol. 40, pp. 582-596.

Roland Berger (2002) Full-Shelf Satisfaction. Reducing out-of-stocks in the grocery channel. Grocery Manufacturers of America (GMA).

Roland Berger (2003) ECR-Optimal Shelf Availability. ECR Europe

Salzberg, S., (1997) On comparing classifiers: pitfalls to avoid and a recommended approach, Data Mining and Knowledge Discovery, 1, 317-328

Urban, T. L. (1998) An inventory-theoretic approach to product assortment and shelf-space allocation. Journal of Retailing, Vol. 74, No 1, pp 15–35.

Van Rijsbergen, C. (1979). *Information Retrieval.* 2nd edition, London, Butterworths

Vuyk, C. (2003) Out-of-Stocks: A Nightmare for Retailer and Supplier. Beverage World, 122, 2, 55. Available at: www.beverageworld.com

Yang, M. (2001) An efficient algorithm to allocate shelf space. European Journal of Operational Research, 131, pp. 107-118.