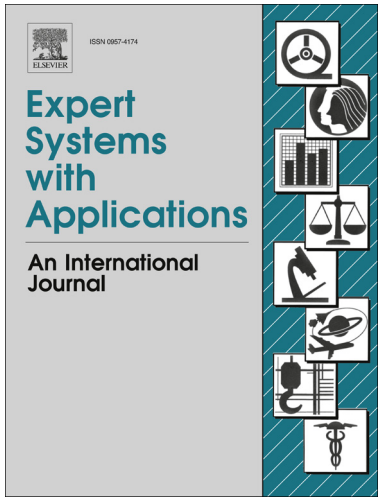

A Decision-Making Framework for Precision Marketing

Zhen You, Yain-Whar Si, Defu Zhang, XiangXiang Zeng, Stephen C.H. Leung, Tao Li

PII: S0957-4174(14)00800-8
DOI: http://dx.doi.org/10.1016/j.eswa.2014.12.022
Reference: ESWA 9746

To appear in: *Expert Systems with Applications*

Please cite this article as: You, Z., Si, Y-W., Zhang, D., Zeng, X., Leung, S.C.H., Li, T., A Decision-Making Framework for Precision Marketing, *Expert Systems with Applications* (2014), doi: http://dx.doi.org/10.1016/j.eswa.2014.12.022

A Decision-Making Framework for Precision Marketing

Zhen You[1], Yain-Whar Si[2], Defu Zhang[1,*], XiangXiang Zeng[1], Stephen C.H. Leung[3], Tao Li[1,4]

[1] Department of Computer Science, Xiamen University, Xiamen 361005, China.

[2] Department of Computer and Information Science, University of Macau, Macau

[3]Department of Management Sciences, City University of Hong Kong, Hong Kong

[4]School of Computer Science, Florida International University, Miami, FL, USA

Emails: uzhen@foxmail.com, fstasp@umac.mo, {dfzhang,xzeng}@xmu.edu.cn, mssleung@cityu.edu.hk, taoli@cs.fiu.edu

Abstract

Precision marketing offers personalized customer service and is used to help enterprises increase their profits by means of high-efficiency marketing. This paper presents a novel decision-making framework for precision marking using data-mining techniques. First, this study presents a trend model to accurately predict monthly supply quantity; second, it uses a RFM (Recency, Frequency and Monetary) model to select attributes to cluster customers into different groups; third, it uses CHAID decision trees and Pareto values to identify important attribute values to distinguish different customer groups; and finally, it creates different supply strategies targeting each customer group. The objective of the proposed precision-making framework is to help managers identify the potential characteristics of different customer categories and put forward appropriate precision marketing strategies, which can greatly reduce inventory for every customer category. The real-world data from a company in China were collected and used in a case study to illustrate how to implement the proposed framework. This case study demonstrates that our proposed decision-making framework is efficient and capable of providing a very good precision marketing strategy for enterprises.

Keywords: Data mining; Decision tree; Forecasting; Precision marketing; Decision-Making;

1. Introduction

Due to the accelerated pace of economic globalization and increasing market competition, economic pressures and competition have led enterprise managers to face the problem of choosing the right strategic decision-making policies for selling the right products to the right customers at the right time, such that the companies can increase their profits. Recently, it has been recognized that precision marketing has become a key means of generating profit and is becoming increasingly important as customers become better informed about the products and their rights as consumers. The availability of customer data and transaction records provides better understanding of customers' consumption behaviours and preferences. In the increasingly competitive environment, enterprises have to create a decision-making

* Corresponding author. Tel.: +86-18959217108; Fax: +86-0592-2580258.
E-mail: dfzhang@xmu.edu.cn

model for precision marketing that provides appropriate strategies to manage the market positioning system for fulfilling their customers' needs. The research motivation of this paper is stemmed from a real-world project**.** This project considers a marketing problem where the supplier or manufacturer provides different products for retail customers, of which some may sell well in some customer segments and some may not. Products that are not sold will be returned back to the supplier. Therefore, the supplier needs to find a good marketing strategy that minimizes goods in stock and satisfies the supplier, retailers and consumers. In recent years, the decision-making problems have received much attention due to a wide range of real-world applications. Many decision-making techniques have been proposed in literature. Chen & Wang (2009) presented multi-criteria optimization and compromise solutions. Saen (2010) developed a technique for order performance via similarity to ideal solution. Hsu et al. (2010) developed a nonlinear programming for decision-making, and Lin et al. (2011) presented a Linear programming for decision making.

When discussing the decision making in this context, it is important to pay attention to the role of artificial intelligence (AI). The decision support framework of AI is considered to be a major tool for obtaining information related to historical data collections using some artificial Intelligence technique such as genetic algorithm (GA), artificial colony optimization (Ghasab et al., 2014), and other data-mining techniques (Chai et al., 2013). Nowadays, data-mining, which can extract useful customer information and discover the hidden customer's behaviours from big data, has a great influence on guiding decision making and forecasting the effects of decisions. Guo et al. (2009) used a hierarchical potential support vector machine (HPSVM) for supplier selection and improved the accuracy. Chang & Hung (2010) presented a rough set theory (RST) to analyze the rules of supplier selection and guide the decision making. Chai et al. (2013) conducted a systematic review of literature about the application of decision-making techniques in supplier selection. They classified these techniques into three categories: Multi-criteria decision making techniques (MCDM), mathematical programming techniques (MP), and artificial intelligence techniques (AI). MCDM is a methodological framework that aims to provide decision makers a knowledgeable recommendation amid a finite set of alternatives (Chai et al., 2013). Wang et al. (2014) developed two kinds of prioritized aggregation operators of IVHFLNs for MCDM.

However, in practice, it is not an easy task to choose a good data mining tool since, each data-mining tool has its own advantages and disadvantages. For example, artificial neural networks (ANNs) involve too many hidden neurons and training parameters (Zhang et al., 2005). Its disadvantages include the "black box" nature, greater computational burden and proneness to over-fitting. However, an advantage of ANN is that ANN has the ability to implicitly detect complex nonlinear relationships between dependent and independent variables. Decision trees are simple to use and easy to understand and they offer many advantages compared with other decision-making tools. However, the disadvantages of decision trees include their instability and relatively low performance. Hence, there is no single best model for all the cases (Akın, 2015). Recently, researchers attempt to combine different models together by considering their respective advantages. Guo et al. (2013) presented a multivariate intelligent decision-making mode which combined

three different model to complete every phase of the decision making process. Tadić et al. (2014) developed a novel hybrid MCDM model that combines fuzzy decision making trial and evaluation laboratory model to provide support to decision makers. Furthermore, Yan & Ma (2015) proposed a novel two-stage group decision-making approach to uncertain quality function deployment.

To the best of our knowledge, the marketing problem considered in this paper is still a new problem. The objective of this paper is to propose a decision-making framework for combining various data-mining algorithms to achieve precision marketing of real products. Real-world data that include historical monthly supply (quantity) and information of every customer were collected from a company in China. The goal is to find a model that can classify targeted customers and predict supply quantity and then provide a strategy for precision marketing, i.e., deciding the quantity of products that every store needs. Depending on the different characteristics and requirements of each phase of the marketing model, the decision-making framework uses four data-mining models/algorithms, which are K-means algorithm, decision tree, Pareto ratio method, and RFM Model, for decision-making. Overall, the purpose of this study is to generate, using data-mining techniques, a decision-making model for products' precision marketing. The proposed decision-making framework is more accurate due to its integrated precision strategies which combine prediction models, clustering, and classifying model. Moreover, the proposed framework incorporates RFM model and Pareto ratio into the process of customer segmentation so that the generated strategies are more convincing.

The rest of this paper is organized as follows. Section 2 introduces related works on common data-mining models and algorithms. Section 3 describes the methodology in our proposed framework briefly. Section 4 presents a case study and results of analyses. Finally, Section 5 concludes the paper.

2. Related Works

Since a single data-mining model may only be suitable for a specific problem, such as prediction or clustering or classification, in the proposed framework, we have combined four data-mining models or algorithms to derive a precision marketing strategy for enterprises. The literature on these four models or algorithms was reviewed below.

2.1. K-means Algorithm

Clustering is the process of grouping a set of physical or abstract items into classes of similar items where the groups are either meaningful or useful, or both. A well-known clustering algorithm is K-means, which was first proposed by MacQueen (1967). The accuracy of this algorithm depends on the initialization and the number of clusters (Mesforoush & Tarokh, 2013). The basic idea of K-means is to discover k clusters, such that the records within each cluster are similar to each other and distinct from the records in other clusters. K-means is an iterative algorithm: an initial set of clusters is defined and the clusters are repeatedly updated until no further improvement is possible (or the number of iterations exceeds a specified threshold). The K-means algorithm is widely used to pre-process data or for clustering because of its simplicity and efficiency (Mesforoush & Tarokh, 2013). K-means has been widely used to effectively identify the valuable customers and develop the related

marketing strategies (Wei et al., 2013; Mesforush & Tarokh, 2013). In particular, Cheng & Chen (2009) use the RFM model and K-means to perform customer relationship management, and the experimental results demonstrate that their proposed model is an effective method in customer value analysis.

2.2. Decision Tree

A decision tree is an efficient data mining algorithm with a strong explanatory capability (Zhang et al., 2010). This study uses the Chi-squared Automatic Interaction Detector (CHAID) decision tree, a decision tree algorithm that is a type of database segmentation. The concept of CHAID was first published in (Kass, 1980). CHAID is used for prediction and classification. Like other decision trees, CHAID's advantages are that its output is highly visual and easy to interpret. A recent study indicates that CHAID is superior to judgment RFM for identification of likely responders (McCarty & Hastak, 2007). CHAID is similar to the RFM approach because it can identify the terminal nodes that will break even with respect to expected profit and costs. Therefore, CHAID was widely adopted used for segmenting customers and extracting rules that show the associations between the input and output variables (McCarty & Hastak, 2007; Chen & Wang, 2009; Mistikoglu et al., 2014). In addition, Coussement et al. (2014) demonstrate that, when performing customer segmentation at different levels of uncertainty, CHAID outperforms RFM and logistic regression. Furthermore, CHAID can also assess its robustness to data accuracy problems. So in our study, we use CHAID to segment customers.

2.3. Pareto Ratio

The concept of Pareto ratio is from the "Pareto Principle" (also known as the 80-20 rule, the law of the vital few or the principle of factor sparsity) which states that for many events, roughly 80% of the effects come from 20% of the causes. This is a common rule of thumb in business. This distribution is relevant to entrepreneurs and business managers, stating, for example, that 80% of the profits come from 20% of the customers (McCarty & Hastak, 2007). This study uses the ordered percentage of marketing / the proportion of customers (i.e., the Pareto Ratio) to identify customers by category (large or small). A higher ratio indicates a larger category of customers. Our empirical study shows that the Pareto Ratio performs quite well in the experiments. Our future work is to investigate other alternative choices for customer identification.

2.4. RFM Model

The RFM model was proposed by Hughes (2000). This model is popular in customer value analysis and has been widely used in measuring customer lifetime value (Cheng & Chen, 2009) and in customer segmentation and behavior analysis (Chen et al. 2012). The RFM model also has been used in several cases, especially in choosing clustering indexes. Recently, researchers have used RFM attributes and the K-means method to improve customer relationship management (CRM) for enterprises (Cheng & Chen, 2009; Wei et al., 2013). In this study, we use the RFM model to choose the variable for clustering so that the clustering standards are established objectively.

The RFM segmentation model is a model that differentiates important customers according to three variables: customers' consumption interval, frequency and

amount of money.

1) R represents “recency”, which is defined as the interval between the time of the latest consuming behavior and the present; the shorter the interval, the greater the value of R.
2) F represents “frequency”, which is defined as the frequency of consuming behavior over a period of time.
3) M represents “monetary”, which is defined as money value of consumption over a period of time.

Research shows that the greater the value of R and F, the more likely the corresponding customers are to have more new trade with enterprises. Moreover, the greater the value of M is, the more likely the corresponding customers are to respond to an enterprise’s products and services again.

The RFM method is very effective at customer segmentation. It sorts customers by their consuming data first, positioning the most recent customer at the front. This classifies the customers into several groups. Then, F and M are standardized and sorted in the same way as described above. Each customer is then positioned in a three-dimensional space, corresponding to a coordinate of R, F and M. With these RFMs sorted in descending order, the groups of customers are classified proportionately.

3. Decision-making Framework

This section details the method employed in this study. The decision-making framework that decides the weekly supply quantity of product among customers includes the following four components (Fig. 1): (1) Data preparation: prepare and analyze the quality of customers to determine their objectives analytically; (2) Supply quantity forecasting: predict the monthly supply quantity by using the appropriate model; (3) Customer classification: analyze every dimension of the customers’ information using a RFM model, clustering algorithm, decision trees algorithm and Pareto ratio to classify the customers’ objectives; and (4) Marketing decision-making: make appropriate supply decisions based on the characteristics of the various customer categories.

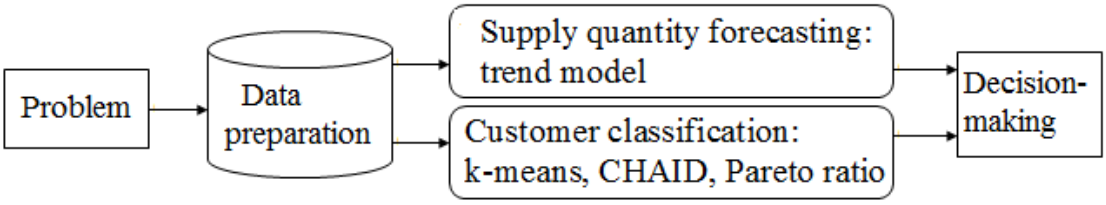

Fig.1 A decision-making framework.

3.1 Data preparation

In order to classify the customers, this paper focuses on customers who have previously ordered products. Therefore, we pre-process the data before building a decision-making model. First, the data of customers, whose ordered quantity or ordered times is zero, is discarded. Next, we select the attributes which would be

used in the data mining procedure of this research. Based on RFM analysis, input attributes to be used for clustering the customers are chosen.

3.2 Supply quantity forecasting

There are many data mining algorithms such as SVM, neural network and ARIMA that can be used to predict monthly supply quantity. However, our experimental study shows that these algorithms are not suitable to predict monthly supply quantity with seasonal characteristics because the supply data for each product are not large enough for learning and testing.

In this subsection, we present a trend model to predict the monthly supply quantity. The model, based on the historical trend of supply data, adjusts the predicted supply quantity. This trend is indicated by the change of mean value. For example, to predict the *j-th* month's supply quantity of *m-th* year, the method is as follows:

1) calculate the mean value of the *j-th* month's supply quantity, using the last three years' supply data, for the current year *m*. We set the mean value as *A[j]*, which is,

$$A[j] = \frac{S[m-3, j] + S[m-2, j] + S[m-1, j]}{3}$$

where $S[m, j]$ denote the supply quantity of *j-th* month in *m-th* year.

2) calculate the mean value of the *(j-1)th* month's supply quantity using the last three years' supply data, namely, calculate *A[j-1]* by Step (1).

3) calculate the predicted supply quantity = last month's supply quantity + *(A[j]-A[j-1])*, which is

$$F[m, j] = F[m, j-1] + A[j] - A[j-1].$$

where $F[m, j]$ denote the predicted supply quantity of *j-th* month in *m-th* year. It is noted that this model can continuously predict several months' value.

The above model considers the relation between historical data of the same month across different years and has a good forecasting performance. The detailed comparisons between different prediction models are presented in the next section to verify the performance of the proposed model.

3.3 Customer classification

Customer classification is performed using K-means, CHAID and Pareto ratio, as shown in Fig. 1. The K-means algorithm is used to cluster retail customers using the attributes chosen based on RFM analysis. After clustering the customers, the clustering results provide decision attributes for the CHAID algorithm. For the results of CHAID algorithm, if the number of customers in a customer category is very large and there are some high-quality customers within that customer category, then we further divide this customer category into more subcategories to optimize the classification performance.

First, we chose the attribute which influences the business condition of the customers in this customer category as a stable attribute. Then, we used this attribute to subdivide the customers in this customer category by calculating the Pareto ratio in every dimension of this attribute. At last, we subdivided these customers by using the Pareto ratio, in every dimension of this attribute, to contrast

this customer category's Pareto mean value (the mean value of all the Pareto ratios in all dimensions of this attribute), where Pareto ratio is calculated as follows:

$$\text{Pareto ratio} = \frac{\text{the percentage of marketing}}{\text{the percentage of customers}}.$$

3.4 Marketing decision-making

Depending on the customer classification results, the strategy of weekly supply quantity for every customer category could be different. For instance, if the key customers' cumulative ordered quantity span is large and scattered, then the strategy for the key customers' category is to satisfy these customers' demand as much as possible. However, this strategy should not be applied to other categories of customers since their demand could be smaller compared to key customers. Therefore, supply quantity for the customers should be decided based on their cumulative ordered quantity. Details of steps involved in derivation of the strategy for these customers are as follows:

1) determine which attributes are referenced attributes in this strategy.
2) find out the customers whose attribute values satisfy the specific referenced attribute values of supplying target, and calculate the sum of these customers and their order quantity. The sum and their order quantity are denoted as *S* and *OQ*, respectively.
3) calculate intersecting supply quantity percentage (*SQP*) as follows:

$$SQP = \frac{OQ}{SOQ},$$

where *SOQ* denotes all the order quantity of customers in the requested customer category.

4) calculate the intersecting supply quantity (*IQ*) as follows:

$$IQ = \frac{SQ * SQP}{S},$$

where *SQ* denotes the supply quantity of the requested customer category.

5) The strategic supply quantity equals the integer part of *IQ* plus one.

4. Experimental Results

Real-world data used in this section were collected from a company in China. The data include six products: Smith, Howard, David, Edward, Yun, and Edward1. The production information such as monthly supply quantity, ordered quantity, ordered times, D-value, Lever of customers, commercial activities, scale of operations, commercial environment, and marketing type are known. The details are given in Table 1. The data are processed to ensure confidentiality of the company's business information.

Table 1. The attribute information on products.

Attribute	Description	Range
Ordered Quantity	Cumulative ordered quantity purchased by the customers during the period from January to June (inclusive)	Numeric
Ordered Times	Cumulative ordered times in the period from January to June (inclusive)	Numeric
D-value	D-value of week during which customers made their final purchase, prior to and including July	Numeric
Level of customers	The level of the customers' operation	1,2,3,4,5
Commercial activities	Type of commercial activities of customers, such as "supermarket", "specialty store", etc.	A…..Q is used to represent every type.
Scale of operation	Scale of customer operation	Large, Middle, Small scale
Commercial Environment	Type of commercial environment of customers, such as "street", "residential area", etc.	a…..k is used to represent every type.
Marketing Type	Location of customer operation	City, Village

4.1 Comparison of prediction models

In order to verify the performance of the trend prediction model mentioned in Section 3, we use Eviews (SARIMA), SPSS (SARIMA) and Support Vector Machine (SVM) to predict different products' supply quantity. The relative error is chosen as the metrics, and is calculated as follows:

$$\text{Error rate} = \frac{|F[m, j] - R[m, j]|}{R[m, j]} * 100\%$$

Where $F[m, j]$ denotes the predicted supply quantity of *j-th* month in the *m-th* year, $R[m, j]$ denotes the real supply quantity of *j-th* month in the *m-th* year. Data of six products (Smith, Howard, David, Edward, Yun and Edward1) from January of 2007 to December of 2010 are used to predict six months' value from January to June in 2011. The relative error rate of different models is shown in Table 2.

Table 2. Error rate of predicting supply

Product name	SPSS(SARIMA)	Eviews(SARIMA)	SVM	Trend model
Smith	24%	53%	21%	13%
Howard	12%	41%	22%	17%
David	70%	33%	17%	13%
Edward	149%	55%	102%	33%
Yun	29%	38%	38%	33%
Edward1	85%	30%	61%	13%
Average	62%	42%	43%	**20%**

The data in Table 2 reveal that the average error rate of the trend model is the lowest among four prediction methods. We find that the proposed model is intuitive

and efficient, and outperforms the existing well-known models. The reason for this may be that Eviews, SPSS and SVM need more historical supply quantity data to build a good prediction model.

4.2 A case study

Without loss of generality, we select Smith product as an example to demonstrate the use of the proposed decision-making framework.

Based on RFM analysis, we assess the importance of customers' satisfaction, ordered frequency and the purchasing time interval. Therefore, the three variables are used as clustering indexes. Moreover, the objective of the CHAID decision trees is to identify activities of customers who have a high ordered quantity of product. The following attributes are used as independent variables: "level of customers", "commercial activities" and "scale of operation". The input attributes for this model are shown in Table 1.

After analysing the data, we discovered that of 24,624 customers, 451 customers never ordered any products; i.e. the cumulative ordered quantities or cumulative ordered times are zero. We ignored these 451 customers and considered only 24,173 customers as objective customers for this case study.

4.2.1 Classification results

From this experiment, we discovered that when we clustered three categories, the distribution of "commercial activity" of key customer is significantly dispersed. On the contrary, when we clustered four or five categories, the key customers' "commercial activity" is concentrated and was clearly identified. The results of clustering four and five categories are shown in Tables 3 and 4, respectively, in which the values are the number of customers.

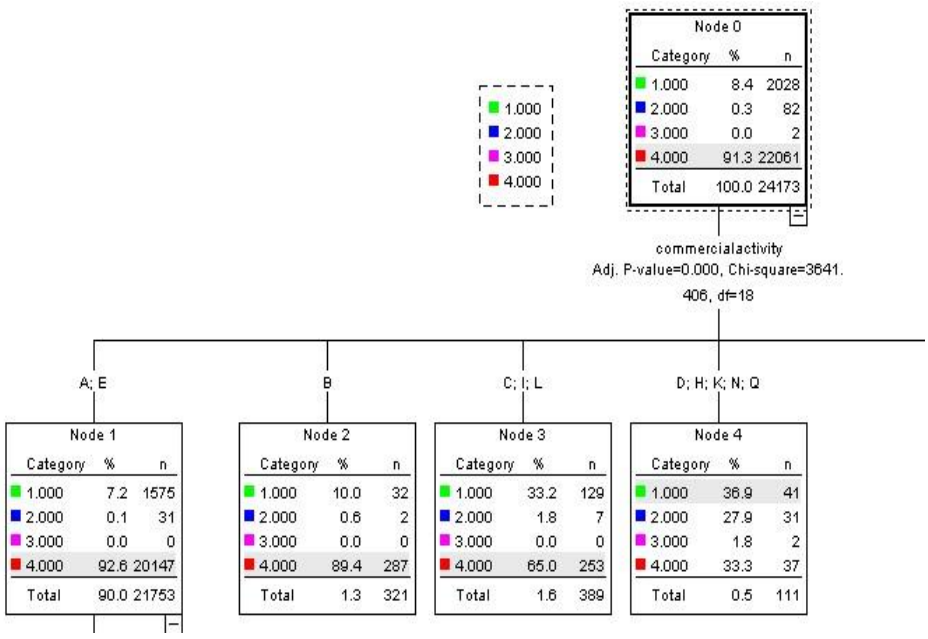

Fig. 2. The results of CHAID decision tree for four categories.

The mean values in Table 3 show that the customers in Categories 2 and 3 are key customers, whereas the customers in Categories 1 and 4 are middle and small (namely, unimportant customers), respectively. To identify the "commercial activity"

of key customers, we use “level of customers”, “commercial activity” and “scale of operation” as input attributes for the CHAID decision trees. A part of the resulting chart of decision tree for four categories is shown in Fig. 2.

Based on Fig. 2, we determined that Node 4 contains all customers who belong to Category 3 and almost one-third of the Category 2 customers; these customers’ “commercial activity” is included in the set: D, H, K, N and Q (see table 1). The other customers’ “commercial activity” is significantly dispersed.

Table 3. The results of clustering four categories by k-means.

		Attributes		
		Cumulative ordered quantity	Cumulative ordered times	D-value of week
Category 1	Total	2,028	2,028	2,028
	Mean value	102	19	3
	Max value	61	27	24
	Min value	260	1	0
Category 2	Total	82	82	82
	Mean value	382	20	1
	Max value	897	27	18
	Min value	262	4	0
Category 3	Total	2	2	2
	Mean value	1,472	27	0
	Max value	1,669	27	0
	Min value	1,275	27	0
Category 4	Total	22,061	22,061	22,061
	Mean value	16	7	8
	Max value	64	27	26
	Min value	1	1	0
Overall	Total	24,173	24,173	24,173
	Mean value	24	8	8
	Max value	1,669	27	26
	Min value	1	1	0

From the mean value in Table 4, we determined that customers of Categories 2 and 3 are key customers; customers of Categories 1, 4 and 5 are small customers. We then used the CHAID algorithm to determine the key customers’ commercial activities. A part of the resulting chart of decision tree for five categories is shown in Fig. 3.

Table 4. The results of clustering five categories by k-means.

		Attributes		
		Cumulative quantity of ordering	Cumulative times of ordering	D-value of weeks
Category 1	Total	120	120	120
	Mean value	331	21	2
	Max value	705	27	0
	Min value	223	4	18
Category 2	Total	1	1	1
	Mean value	1,669	27	0
	Max value	1,669	27	0
	Min value	1,669	27	0
Category 3	Total	2	2	2
	Mean value	1,007	22	0
	Max value	1,275	27	0
	Min value	879	16	0
Category 4	Total	21,616	21,616	21,616
	Mean value	15	7	8
	Max value	58	27	26
	Min value	1	1	0
Category 5	Total	2,435	2,435	2,435
	Mean value	92	18	3
	Max value	222	27	24
	Min value	54	1	0
Overall	Total	24,173	24,173	24,173
	Mean value	24	8	8
	Max value	1,669	27	26
	Min value	1	1	0

According to Fig. 3, we found that Node 4 (D, H, K, N, Q) contains all customers who belong to Categories 2 and 3. Their “commercial activity” is included in the set: D, H, K, N and Q. Other customers’ “commercial activity” is significantly dispersed.

From the analytical results given above, we ensure that the first category of customers (key customer) “commercial activity” is D, H, K, N and Q. This category has 111 customers who account for 0.46% of all customers. Their ordered quantity accounts for 3.9% of the total.

Removing these 111 customers, the remaining 24,062 customers are classified by Pareto ratio because the distribution of their “commercial activity” is significantly dispersed. Therefore, we take advantage of the Pareto ratio to analyse the remaining customers’ “commercial activity”. The results of the analysis are shown in Table 5.

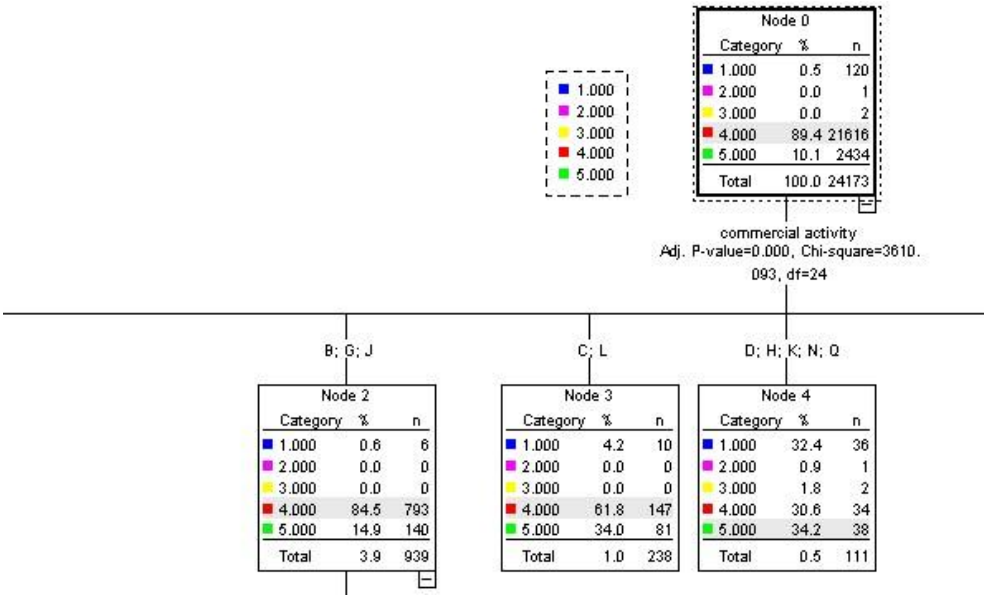

Fig. 3. The result of CHAID for five categories.

Table 5. The Pareto ratio of every commercial activity.

Commercial activity	Customer quantity	Ordered quantity	Customer percentage	Quantity percentage	Pareto value
A	21,445	89.1239%	471,620	83.4442%	0.936272
B	321	1.3341%	8,274	1.4639%	1.097353
C	126	0.5236%	6,981	1.2352%	2.358755
E	308	1.28%	5,712	1.0106%	0.789538
F	887	3.6863%	35,588	6.2966%	1.708109
G	496	2.0613%	15,147	2.68%	1.300112
I	151	0.6275%	7,876	1.3935%	2.22057
J	122	0.507%	4,407	0.7797%	1.537868
L	112	0.4655%	8,457	1.4963%	3.214652
M	66	0.2743%	893	0.158%	0.576028
O	22	0.0914%	140	0.0248%	0.27092
P	6	0.0249%	97	0.0172%	0.688266
Total	24062	100%	565,192	100.0%	-

From Table 5, we determine that the customers' quantity of commercial activity "A" accounts for 88.6% of the remaining customers, but its ordered quantity is only 80.2% of the remaining ordered quantity. Thus, "A" is typical of small customers. We classified the third category of customers whose Pareto ratio is not greater than 1.0 and classified the second category of customers whose Pareto ratio is greater than 1.0. The detailed results are summarized in Table 6.

Table 6. The results of the customer classification.

Case Summary						
Category	Commercial activity	Customer quantity	Customer percentage	Order quantity	Quantity percentage	Pareto ratio
1	D	45	0.19%	12,676	2.16%	11.57954
	H	11	0.05%	1,420	0.24%	5.306609
	K	41	0.17%	5,628	0.96%	5.642762
	N	13	0.05%	2,798	0.48%	8.847607
	Q	1	0.00%	329	0.06%	13.52438
	Total	111	0.46%	22,851	3.90%	44.9009
2	B	321	1.33%	8,274	1.41%	1.059576
	C	126	0.52%	6,981	1.19%	2.277553
	F	887	3.67%	35,588	6.05%	1.649307
	G	496	2.05%	15,147	2.58%	1.255355
	I	151	0.62%	7,876	1.34%	2.144126
	J	122	0.50%	4,407	0.75%	1.484926
	L	112	0.46%	8,457	1.44%	3.103986
	Total	2,215	9.15%	86,730	14.76%	12.97483
3	A	21,445	88.71%	471,620	80.20%	0.90404
	E	308	1.27%	5,712	0.97%	0.762358
	M	66	0.27%	893	0.15%	0.556197
	O	22	0.09%	140	0.02%	0.261593
	P	6	0.02%	97	0.02%	0.664572
	Total	21,847	90.36%	478,462	81.36%	3.14876
Total	A	21,445	88.71%	471,620	80.20%	0.90404
	B	321	1.33%	8,274	1.41%	1.059576
	C	126	0.52%	6,981	1.19%	2.277553
	D	45	0.19%	12,676	2.16%	11.57954
	E	308	1.27%	5,712	0.97%	0.762358
	F	887	3.67%	35,588	6.05%	1.649307
	G	496	2.05%	15,147	2.58%	1.255355
	H	11	0.05%	1,420	0.24%	5.306609
	I	151	0.62%	7,876	1.34%	2.144126
	J	122	0.50%	4,407	0.75%	1.484926
	K	41	0.17%	5,628	0.96%	5.642762
	L	112	0.46%	8,457	1.44%	3.103986
	M	66	0.27%	893	0.15%	0.556197
	N	13	0.05%	2,798	0.48%	8.847607
	O	22	0.09%	140	0.02%	0.261593
	P	6	0.02%	97	0.02%	0.664572
	Q	1	0.004%	329	0.06%	13.52438
	Total	24,173	100.0%	588,043	100.0%	-

Table 6 shows that the 111 customers in the first category account for 0.60% of all customers but the ordered quantity is 3.89% of the entire ordered quantity. The Pareto ratio of every “commercial activity” in the first category is very high; thus,

customers in the first category are “key customers”. The quantity of the second category, with 2,309 customers, accounts for 9.55% of the total quantity whereas the ordered quantity is 14.90% of the entire ordered quantity, and the Pareto ratio of every commercial activity in the second category is somewhat high. Thus, customers of the second category are “demonstrate customers” (the customers whose ordered quantity percentage is little higher than customers quantity percentage). The 21,847 customers in the third category account for 89.9% of all customers but the ordered quantity is only 81.3% of the entire quantity and the Pareto ratio of every commercial activity in the third category is small; thus, the customers of the third category are “small customers”.

We classified the customers into three categories, but the number of customers in the third category is too large. The third category, with 21,847 customers, includes 89.9% of the total number of customers. To identify high-quality customers within the third category, we optimized the classification of the third customer category.

It should be noted that the attribute “Commercial Environment” influences the business condition of “small customers” and is considered a stable attribute. Thus, we used “Commercial Environment” to subdivide the “small customers”. As shown in Table 7, when we placed the customers whose Pareto ratio is greater than the third category’s Pareto mean value into the third category and placed the others into the fourth category. We obtained the results as shown in Table 7.

Table 7. The results of classification optimisation.

Commercial Environment	Customer quantity	Customer percentage	Order quantity	Quantity percentage	Pareto value	Category
A	2,917	13.35%	63,779	13.33%	0.998356	4
B	3,302	15.11%	65,420	13.67%	0.904644	4
C	6,158	28.19%	126,882	26.52%	0.940816	4
D	347	1.59%	8,765	1.83%	1.153365	3
E	5,948	27.22%	137,508	28.74%	1.055605	3
F	2,558	11.71%	60,427	12.63%	1.078636	3
G	335	1.53%	8,142	1.70%	1.109764	3
H	261	1.19%	7,201	1.51%	1.259786	3
I	4	0.02%	87	0.02%	0.993124	4
J	2	0.01%	38	0.01%	0.867557	4
K	15	0.07%	213	0.04%	0.648385	4
Total	21,847	1	478,462	1	11.01004	-

Then, we calculated the Pareto mean value as 1.00913 and subdivided the third and fourth categories according to the terms mentioned above. After segmenting the customers, we obtained the results as shown in Table 8.

Taking advantage of the “commercial environment”, we subdivided the “small customers” into the third and fourth categories. The third category, with 9449 customers, accounted for 39% of all customers and its Pareto ratio is 0.9743. The fourth category, with 12,398 customers, accounts for 51.3% of all customers and its Pareto ratio is 0.85.

In summary, we obtained the structure for four customers’ categories as shown in Fig. 4(a)-(f).

Table 8. The results of classification.

Category	Customer quantity	Customer percentage	Order quantity	Quantity percentage	Pareto value
1	111	0.46%	22,851	3.90%	44.9009
2	2,215	9.15%	86,730	14.76%	12.97483
3	9,449	39%	222,043	38%	0.9743
4	12,398	51.3%	256,419	43.6%	0.85
Total	24,173	100.0%	588,043	100.0%	59.7

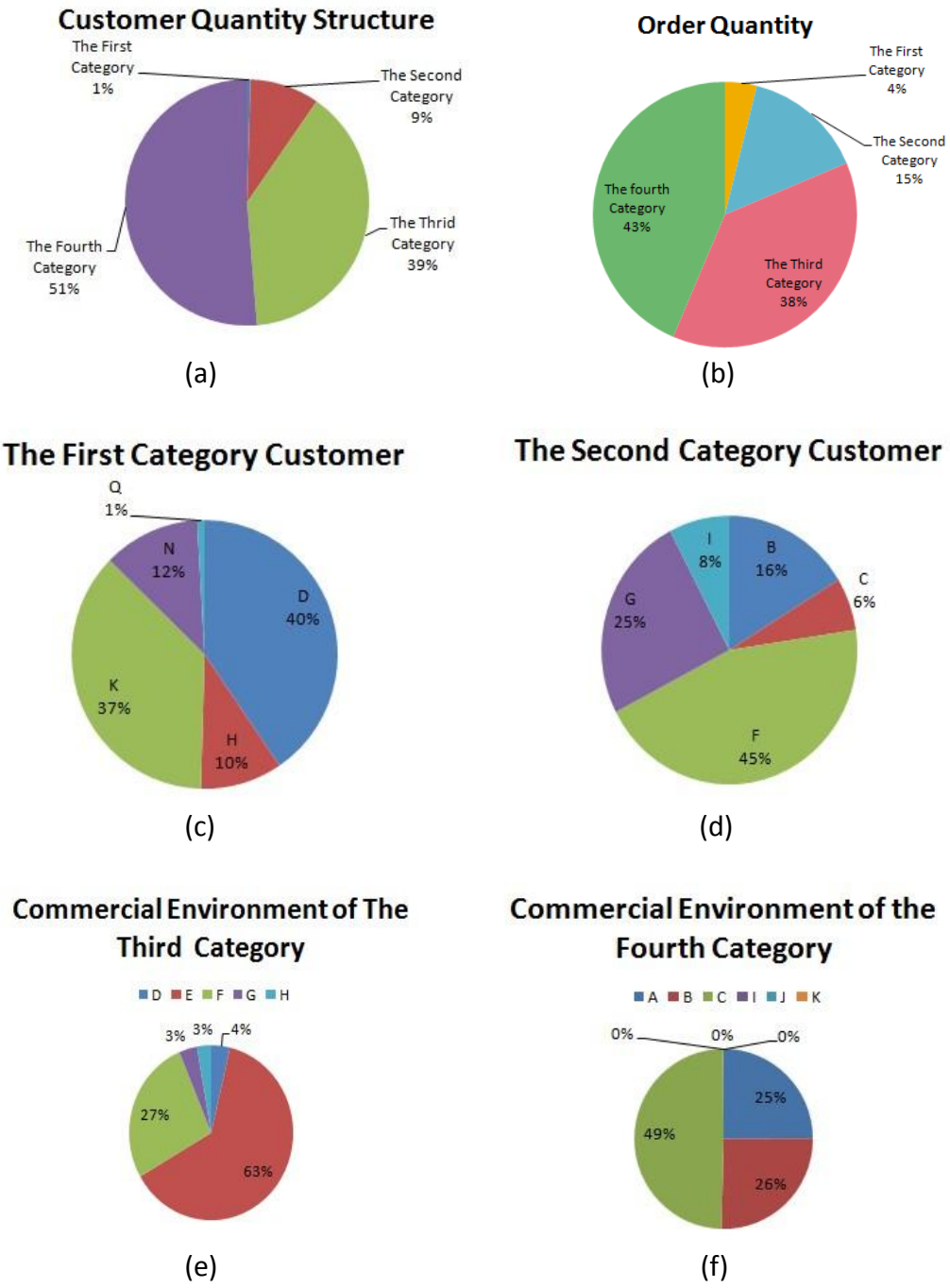

Fig. 4. The structure of the four customer categories

4.2.2 Decision on the weekly supply quantity

By using the trend prediction model given above, we calculated the supply quantity of "Smith" product for the seventh month as 152,299. If customers order products every working day, then we should supply 7252 (152,299/21 = 7252) cartons of this product every working day or 36,260 cartons of this product every five-day working week. As shown in Table 9, we then calculated the supply quantity of every category in terms of the four categories' ordered percentage (Table 9).

Table 9. The results of weekly supply quantity.

Customer Category	Customer quantity	Ordered quantity	Customer percentage	Quantity percentage	Weekly supply quantity
1	111	22,851	0.46%	3.90%	1,414
2	2,215	86,730	9.15%	14.76%	5,366
3	9,449	222,043	39%	38%	13,779
4	12,398	256,419	51.3%	43.6%	15,701
Total	24,173	588,043	100.0%	100.0%	36,260

(1) The supply decision for the first customer category

According to our model, we can obtain the ordered quantity frequency for the first customer category as shown in Fig. 5 where we can see that the ordered quantity span is large and scattered. A unified strategy will not meet the total demand and the business will lose some of its customers. Therefore, a better strategy would be to satisfy the customers' demand as much as possible. However, the total supply quantity cannot exceed the supply quantity allotted for the entire month. Thus, "Smith" product should supply as much as possible to meet the first category customers' demand, but the total quantity cannot exceed 1414.

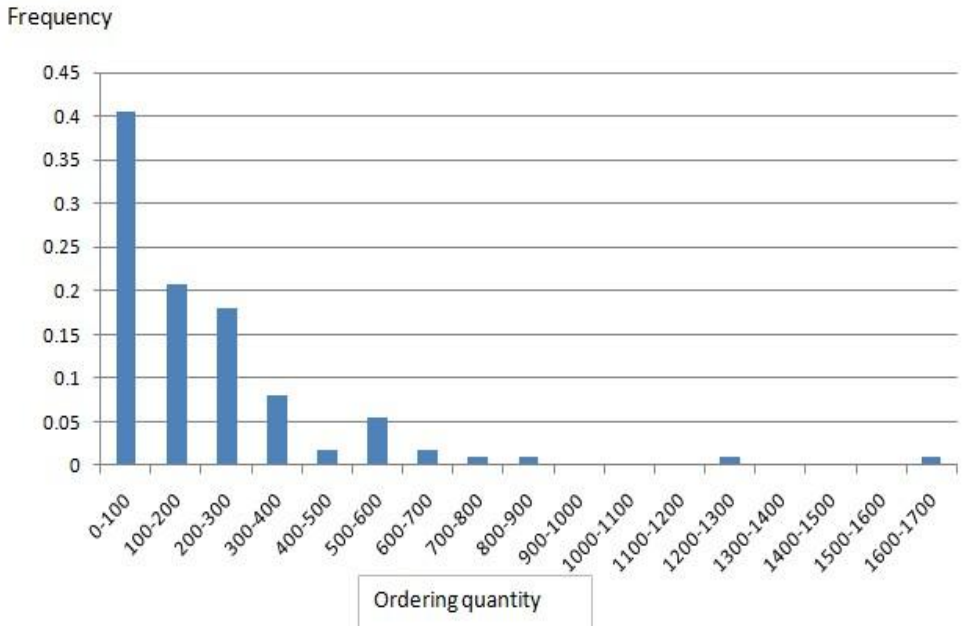

Fig. 5. The ordered quantity frequency for the first customer category

Table 10. The initial supply strategy for the second customer category.

Commercial activity	Marketing scale	Marketing type	Customer level				
			5	4	3	2	1
B	Big scale	city	3	2	0	2	0
		village	1	5	1	0	0
	Middle scale	city	3	2	3	2	1
		village	0	1	2	1	1
	Small scale	city	2	2	1	1	1
		village	0	1	1	0	0
C	Big scale	city	5	4	3	5	3
		village	0	2	10	0	0
	Middle scale	city	0	3	5	0	0
		village	0	1	0	0	0
	Small scale	city	0	3	0	0	0
		village	0	0	1	0	0
F	Big scale	city	8	4	1	3	1
		village	4	0	0	0	0
	Middle scale	city	3	2	2	2	2
		village	4	4	0	0	2
	Small scale	city	2	2	1	2	2
		village	3	2	2	0	1
G	Big scale	city	7	4	5	0	3
		village	0	3	0	0	0
	Middle scale	city	3	2	2	1	1
		village	4	4	0	4	0
	Small scale	city	3	1	1	1	1
		village	1	1	1	0	0
I	Big scale	city	5	6	0	0	0
		village	0	0	0	0	0
	Middle scale	city	4	3	1	0	0
		village	5	0	0	0	0
	Small scale	city	3	1	1	3	3
		village	0	0	0	0	0
J	Big scale	city	5	2	3	1	0
		village	0	0	0	0	0
	Middle scale	city	3	2	4	3	2
		village	0	1	1	3	1
	Small scale	city	9	2	0	1	4
		village	0	1	0	0	0
L	Big scale	city	6	3	7	2	0
		village	17	11	9	0	0
	Middle scale	city	3	4	1	1	11
		village	0	0	0	0	0
	Small scale	city	0	0	1	0	0
		village	0	0	0	0	0

(2) The supply decision for the second customer category

Because the customers in the second category are “demonstrate customers”, the strategy which satisfies the customers’ demand may result in unnecessary waste. So, in order to achieve accurate supply, the supply strategy of the second customer category must consider these attributes: “customer activity”, “marketing scale”, “marketing type”, “customer level” and the customers’ quantity intersecting “customer activity”.

According to the model above, we calculated the initial supply quantity by using the method mentioned in Section 3, and then we obtained the initial supply strategy as shown in Table 10.

Table 11. The initial supply strategy for the third customer category.

Commercial environment	Marketing scale	Marketing type	Customer level				
			5	4	3	2	1
D	Big scale	city	4	21	1	0	5
		village	1	1	0	0	0
	Middle scale	city	1	1	1	0	1
		village	1	1	1	0	0
	Small scale	city	0	1	1	1	1
		village	1	1	1	1	0
E	Big scale	city	2	1	1	1	1
		village	2	2	2	0	0
	Middle scale	city	1	1	1	1	1
		village	1	1	1	1	1
	Small scale	city	1	1	1	1	1
		village	1	1	1	1	1
F	Big scale	city	2	1	1	1	0
		village	2	2	2	1	0
	Middle scale	city	1	1	1	1	1
		village	1	1	1	1	1
	Small scale	city	1	1	1	1	1
		village	1	1	1	1	1
G	Big scale	city	1	2	0	0	0
		village	2	2	0	1	0
	Middle scale	city	0	1	0	0	1
		village	2	1	1	1	1
	Small scale	city	0	1	1	0	1
		village	0	1	1	1	0
H	Big scale	city	2	1	0	0	0
		village	3	2	1	3	0
	Middle scale	city	1	1	1	1	1
		village	1	1	2	1	1
	Small scale	city	0	1	1	1	1
		village	1	1	1	1	1

(3) The supply decision for the third customer category

The supply strategy for the third category is identical to the supply strategy in relation to the second category. But the third customer category and the fourth customer category are classified by “commercial environment”, so we took the “commercial environment” into account instead of “commercial activity”. The supply strategy is expressed in Table 11.

Table 12. The initial supply strategy for the fourth customer category.

Commercial environment	Marketing scale	Marketing type	Customer level				
			5	4	3	2	1
A	Big scale	city	3	1	3	2	0
		village	1	1	2	0	1
	Middle scale	city	1	1	1	1	1
		village	1	1	1	1	1
	Small scale	city	1	1	1	1	1
		village	1	1	1	1	1
B	Big scale	city	2	1	1	1	1
		village	2	1	2	0	1
	Middle scale	city	1	1	1	1	1
		village	1	1	1	1	1
	Small scale	city	1	1	1	1	1
		village	1	1	1	1	1
C	Big scale	city	2	1	1	2	1
		village	1	1	3	1	0
	Middle scale	city	1	1	1	1	1
		village	1	1	1	1	1
	Small scale	city	1	1	1	1	1
		village	1	1	1	1	1
I	Big scale	city	1	0	0	0	0
		village	0	0	0	0	0
	Middle scale	city	0	0	1	0	0
		village	0	0	0	0	0
	Small scale	city	0	0	0	0	0
		village	0	0	0	0	0
J	Big scale	city	0	1	0	0	0
		village	0	0	0	0	0
	Middle scale	city	0	0	0	0	0
		village	0	1	0	0	0
	Small scale	city	0	0	0	0	0
		village	0	0	0	0	0
K	Big scale	city	3	1	0	0	0
		village	0	0	0	0	0
	Middle scale	city	1	1	1	0	0
		village	0	0	0	0	0
	Small	city	0	1	0	0	0

	scale	village	0	0	0	0	1

(4) The supply decision for the fourth customer category

The supply strategy for the fourth customer category is identical to the supply strategy for the third customer category, as shown in Table 12.

5. Conclusions

This paper proposes a general decision-making framework which combines a trend prediction model, K-means algorithm, and CHAID algorithm. A trend prediction model is presented to forecast supply quantity, computational results show that the proposed model outperforms the well-known prediction models such as SARIMA and SVM from commercial software. RFM is used to select attributes for k-means algorithm. The result of clustering is used for the decision attribute of CHAID algorithm. Pareto ratio is further used to divide the key customers.

5.1 Research contributions

This proposed framework presented a trend prediction model to forecast the supplying quantity. This developed trend prediction model is more efficient than other predicted models. Moreover, this framework is able to accurately extract the characteristics of customer categories based on the combination of K-means and CHAID decision trees. Comparing with the work of Wei et al. (2013) (which developed the marketing strategies only based on segmenting customers) and the work of Guo et al. (2013) (which proposed a decision-making model only for retail sales forecasting), our proposed framework is more generic, efficient, and adaptive. As far as we know, it is the first decision-making framework for the marketing problem considered in the paper. On this basis, enterprises can make precise supply strategies for different customer categories, which target high potential customers in several ways, including new products and encouraging higher consumption. It is a win-win situation for both parties since enterprises can employ the supply strategies to maximize their profits while maintaining sufficient level of customer satisfaction. In addition, the case study shows that this decision-making framework is efficient and can help enterprises in planning their precision marketing. Our model is applied to a real-world system, and decreases the inventory by about 20%, such that enterprise's profit is improved according to the company's reports.

5.2 Research limitations

Despite these significant contributions, this proposed framework has some limitations. First, the prediction model is simple. Every prediction model has its advantages so that combining them together may achieve better results. Second, the number of clusters in K-means was limited to the results of CHAID trees. Therefore, it's hard to determine the number of clusters. The third potential limitation is the scalability of the prediction model in terms of the size of historical data. We found that the data size greatly influences the accuracy of prediction model.

5.3 Potential future research

Based on the limitation of this paper and the computational results, we propose some potential directions for future research.

The first focused research direction is the prediction model. Our computational results show that every model has its advantages and different models often obtain different results. Therefore, the integrated prediction model, which combines different prediction models used in this paper, will be an interesting research direction.

The second direction for future research is how to select the most suitable model according to real-world applications.

Finally, our experiments show that the accuracy of model is greatly influenced by the choice of parameters. So, developing an algorithm which can dynamically adjust the parameter settings is a good research direction.

Acknowledgment

The work was partially supported by the National Nature Science Foundation of China (61272003) and the Major Program of the National Social Science Foundation of China (Grant no. 13&ZD148).

References

Akın, M. (2015). A novel approach to model selection in tourism demand modeling. *Tourism Management, 48*, 64-72.

Chen, L. Y., & Wang, T. C. (2009). Optimizing partners' choice in IS/IT outsourcing projects: The strategic decision of fuzzy VIKOR. *International Journal of Production Economics, 120*(1), 233-242.

Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications, 36*(3), 4176-4184.

Chang, B., & Hung, H. F. (2010). A study of using RST to create the supplier selection model and decision-making rules. *Expert Systems with Applications, 37*(12), 8284-8295.

Chen, D., Sain, S. L., & Guo, K. (2012). Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. *Journal of Database Marketing & Customer Strategy Management, 19*(3), 197-208.

Chai, J., Liu, J. N., & Ngai, E. W. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications, 40*(10), 3872-3885.

Coussement, K., Van den Bossche, F. A., & De Bock, K. W. (2014). Data accuracy's impact on segmentation performance: Benchmarking RFM analysis, logistic regression, and decision trees. *Journal of Business Research, 67*(1), 2751-2758.

Guo, X., Yuan, Z., & Tian, B. (2009). Supplier selection based on hierarchical potential support vector machine. *Expert Systems with Applications, 36*(3), 6978-6985.

Guo, Z. X., Wong, W. K., & Li, M. (2013). A multivariate intelligent decision-making model for retail sales forecasting. *Decision Support Systems, 55*(1), 247-255.

Ghasab, M. A. J., Khamis, S., Mohammad, F., & Fariman, H. J. (2015). Feature decision-making ant colony optimization system for an automated recognition of plant species. *Expert Systems with Applications, 42*(5), 2361–2370.

Hughes, A. M. (2000). Strategic Database Marketing: the Masterplan for Starting and Managing a Profitable. *Customer-Based Marketing Program*.

Hsu, B. M., Chiang, C. Y., & Shu, M. H. (2010). Supplier selection using fuzzy quality data and their applications to touch screen. *Expert Systems with Applications, 37*(9), 6192-6200.

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied statistics*, 119-127.

Lin, C. T., Chen, C. B., & Ting, Y. C. (2011). An ERP model for supplier selection in electronics industry. *Expert Systems with Applications, 38*(3), 1760-1765.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability,* 1, 281-297.

McCarty, J. A., & Hastak, M. (2007). Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. *Journal of business research, 60*(6), 656-662.

Mesforoush, A., & Tarokh, M. J. (2013). Customer Profitability Segmentation for SMEs Case Study: Network Equipment Company. *International Journal of Research in Industrial Engineering, 2*(1), 30-44.

Mistikoglu, G., Gerek, I. H., Erdis, E., Usmen, P. M., Cakan, H., & Kazan, E. E. (2015). Decision Tree Analysis of Construction Fall Accidents Involving Roofers. *Expert Systems with Applications ,* 42(4), 2256–2263.

Saen, R. F. (2010). Developing a new data envelopment analysis methodology for supplier selection in the presence of both undesirable outputs and imprecise data. *The International Journal of Advanced Manufacturing Technology, 51*(9-12), 1243-1250.

Tadić, S., Zečević, S., & Krstić, M. (2014). A novel hybrid MCDM model based on fuzzy DEMATEL, fuzzy ANP and fuzzy VIKOR for city logistics concept selection. *Expert Systems with Applications, 41*(18), 8112-8128.

Wei, J. T., Lee, M. C., Chen, H. K., & Wu, H. H. (2013). Customer relationship management in the hairdressing industry: An application of data mining techniques. *Expert Systems with Applications, 40*(18), 7513-7518.

Wang, J. Q., Wu, J. T., Wang, J., Zhang, H. Y., & Chen, X. H. (2014). Interval-valued hesitant fuzzy linguistic sets and their applications in multi-criteria decision-making problems. *Information Sciences, 288*, 55-72.

Yan, H. B., & Ma, T. (2015). A group decision-making approach to uncertain quality function deployment based on fuzzy preference relation and fuzzy majority. *European Journal of Operational Research, 241*(3), 815-829.

Zhang, D., Jiang, Q., & Li, X. (2005). A hybrid mining model based on neural network and kernel smoothing technique. In *Computational Science–ICCS, Springer Berlin Heidelberg,* 801-805,

Zhang, D., Zhou, X., Leung, S. C., & Zheng, J. (2010). Vertical bagging decision trees model for credit scoring. *Expert Systems with Applications,37*(12), 7838-7843.

Highlights

- A decision-making framework for precision marking based on data-mining techniques
- A trend model to accurately predict monthly supply quantity
- A RFM (Recency, Frequency and Monetary) model to select customer attributes
- Decision trees and Pareto values are combined for grouping customers
- A real case-study to demonstrate the effectiveness of the proposed framework