DATA MINING APPLICATIONS IN A FORKLIFT DISTRIBUTOR

PRATIWI EKA PUSPITA

T.C.
ULUDAĞ UNIVERSITY
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

DATA MINING APPLICATIONS IN A FORKLIFT DISTRIBUTOR

Pratiwi Eka PUSPITA

Assoc. Prof. Dr. Tülin İNKAYA
(Supervisor)

MASTER OF SCIENCE THESIS
DEPARTMENT OF INDUSTRIAL ENGINEERING

BURSA – 2018

APPROVAL OF THE THESIS

The thesis entitled DATA MINING APPLICATIONS IN A FORKLIFT DISTRIBUTOR which is prepared and submitted by Pratiwi Eka Puspita in partial fulfillment of the requirements for the degree of **Master of Sciences in Department of Industrial Engineering, The Graduate School of Natural and Applied Sciences, Uludağ University** has been examined and approved by the committee,

Supervisor: Assoc. Prof. Dr. Tülin İNKAYA

Head : Prof. Dr. Erdal EMEL
Uludağ University, Faculty of Engineering
Industrial Engineering Dept.,

Member : Assoc. Prof. Dr. Tülin İNKAYA
Uludağ University, Faculty of Engineering
Industrial Engineering Dept.,

Member : Assist. Prof. Dr. Fatma YERLİKAYA-ÖZKURT
Atılım University, Faculty of Engineering
Industrial Engineering Dept.

I accept the above statements

Prof. Dr. Ali BAYRAM
Head of The Graduate School of Natural and Applied Sciences

22 /2 /2018 (Date)

I prepare this document in accordance with the writing rules of the Graduate School of Natural and Applied Sciences, Uludağ University;

- All the information in this thesis has been obtained under the academic rules,
- The visual, audio and written results are delivered in accordance with the ethical conduct,
- As required, the other works are fully cited in accordance with the scientific norms,
- The other works I have referenced are shown,
- The used data is original,
- Any part of this thesis has not been presented as a thesis study in this university or other universities.

I hereby declare that the above information is true.

22/2/2018

Pratiwi Eka PUSPITA

ABSTRACT

Msc. Thesis

DATA MINING APPLICATIONS IN A FORKLIFT DISTRIBUTOR

Pratiwi Eka PUSPITA

Uludağ University
Graduate School of Natural and Applied Sciences
Department of Industrial Engineering

Supervisor: Assoc. Prof. Dr. Tülin İNKAYA

Sales forecasting has a vital role in today's business environment. In a company, accurate and reliable sales forecasting is the fundamental basis for production planning processes. In this study, a data mining-based forecasting methodology is proposed for a forklift distributor. Monthly sales data for 100 different types of forklifts between years 1998 and 2016 are used. The proposed methodology has three stages. In the first stage, items with similar sales patterns are identified using hierarchical clustering. Dynamic time warping (DTW) is used for measuring the similarities among the items. The number of clusters is determined using the heterogeneity and homogeneity criteria. For each cluster, cluster prototypes are found based on cluster medoids and DTW barycenter averaging (DBA) method. In the second stage, features are extracted. In addition to the features that characterize amount, trend, growth, and volatility, new features are proposed to identify the intermittency in the data. Also, the important features are selected using multivariate adaptive regression splines (MARS). Then, support vector regression (SVR) is used as a forecasting model for each cluster prototype. In the final stage, the proposed approach is evaluated according to inventory performance. The numerical analysis shows that the proposed methodology forecasts the sales with reasonable accuracy and low complexity, and provides a reduction in inventory management costs.

Keywords: Data mining, clustering, forecasting, dynamic time warping (DTW), multivariate adaptive regression splines (MARS), support vector regression (SVR)

2018, x + 109 pages.

ÖZET

Yüksek Lisans Tezi

BIR FORKLIFT DAĞITICISINDA VERI MADENCILIĞI UYGULAMASI

Pratiwi Eka PUSPITA

Uludağ Universitesi
Fen Bilimleri Enstitüsü
Endüstri Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Tülin İNKAYA

Satış tahmini bugünün iş ortamında hayati bir role sahiptir. Bir şirkette, doğru ve güvenilir satış tahminleri, üretim planlama sürecinin esas dayanağıdır. Bu çalışmada, bir forklift distribütörü için veri madenciliğine dayalı bir tahmin metodolojisi önerilmiştir. 1998 ve 2016 yılları arasında 100 farklı forkliftin aylık satış verileri kullanılmıştır. Önerilen metodolojinin üç aşaması vardır. İlk aşamada, benzer satış yapıları içeren ürünler hiyerarşik kümeleme kullanılarak belirlenmiştir. Ürünler arasındaki benzerliklerin ölçülmesinde dinamik zaman bükmesi (DTW) kullanılmıştır. Kümelerin sayısı, heterojenlik ve homojenlik kriterleri kullanılarak belirlenmiştir. Her küme için küme prototipleri küme medoidleri ve DTW ağırlık merkezi ortalaması (DBA) metodu temel alınarak bulunmuştur. İkinci aşamada, öznitelikler çıkarılmıştır. Miktar, eğilim, büyüme ve oynaklığı karakterize eden özniteliklerin yanı sıra verideki düzensiz aralıkları belirlemek için yeni öznitelikler önerilmiştir. Ayrıca, önemli öznitelikler çok değişkenli uyarlanabilir regresyon eğrileri (MARS) kullanılarak seçilmiştir. Ardından, her bir küme prototipi için bir tahmin modeli olarak destek vektör regresyonu (SVR) kullanılmıştır. Son aşamada, önerilen yaklaşım envanter performansına göre değerlendirilmiştir. Sayısal analiz, önerilen metodolojinin satışları makul doğruluk ve düşük karmaşıklıkla tahmin ettiğini ve envanter maliyetlerinde azalma sağladığını göstermektedir.

Anahtar Kelimeler: Veri madenciliği, kümeleme, tahmin, dinamik zaman bükmesi (DTW), çok değişkenli uyarlanabilir regresyon eğrileri (MARS), destek vektör regresyonu (SVR)

2018, x + 109 sayfa.i

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisors, Assoc. Prof. Dr. Tülin İNKAYA and Assist. Prof. Dr. Mehmet AKANSEL for the guidance during the research. They are the best supervisors I have ever met. I really benefit their excellent knowledge that is helpful to improve my skill in doing research.

I also would like to thank Prof. Dr. Erdal EMEL, Prof. Dr. İsmail EFİL, Prof. Dr. Cenk ÖZMUTLU, Prof. Dr. Seda ÖZMUTLU, Assoc. Prof. Dr. Ali Yurdun ORBAK, Assoc. Prof. Dr. Fatih CAVDUR, Assoc. Prof. Dr. Betul YAĞMAHAN, Assist. Prof. Dr. Türker ÖZALP, and Assist. Research Dr. İlker KÜÇÜKOĞLU, who let me experience the awesome education in Turkey.

A bunch of thanks goes to my friends, Sara, Hande, İlknur, Muge, Zeynep, Elif, Enis, and many others. Without them as good friends, I couldn't be able to deal with the barrier language during the courses.

Of course, the highest gratefulness to Allah, then to the members of my family. Thanks to my lovely husband, my parents, also sisters and brother for supporting and encouraging me.

TABLE OF CONTENTS

Page

LIST OF NOTATIONS AND ABBREVIATIONS

Notations	**Description**
b	bias
H	cluster prototype
a	coefficient
d	distance
K	Kernel function
α	Lagrangian multiplier
M	number of basis function
k	number of clusters
p	probability belong to a specified class
S	sequence of time-series data
ξ	slack variable
C	total number of classes
s	vector of time-series data
w	warping path
z	weight vector

Abbreviations	**Description**
ANFIS	Adaptive Network-Based Fuzzy Inference System
ARIMA	Autoregressive Integrated Moving-Average
ARMAX	Autoregressive Moving Average Exogenous
BPN	Backpropagation Neural Network
CART	Classification and Regression Tree
CMACNN	Cerebellar Model Articulation Controller Neural Network
CWRT	Cross-Words Reference Template
DBA	DTW Barycenter Averaging
DMF	Data Mining-Based Forecasting
DTW	Dynamic Time Warping
EOQ	Economic Order Quantity
ES	Exponential Smoothing
EWMA	Exponentially Weighted Moving Average
GA	Genetic Algorithm
HC	Hierarchical Clustering
HW	Holt-Winters
ICA	Independent Component Analysis
ID3	Iterative Dichotomiser
LRSVM	Hybridization of Logistic Regression and SVR
MA	Moving Average
MAD	Mean Absolute Deviation

Abbreviations	Description
MAPE	Mean Absolute Percentage Error
MARS	Multivariable Adaptive Regression Splines
MSE	Mean Square Error
NLAAF	Nonlinear Alignment and Averaging Filter
PNN	Probabilistic Neural Network
PSA	Prioritised Shape Averaging
PSO	Particle Swarm Optimization Algorithm
RBF	Radial Basis Function
RFID	Radio Frequency IDentification
RMSE	Root Mean Square Error
RTW	Regression Time Warping
SD	Standard Deviation
ShapeDTW	Shape DTW
STW	Segment-wise Time Warping
SVM	Support Vector Machine
SVR	Support Vector Regression
SWM	Scaled and Warped Matching
WDTW	Weight DTW
WGSS	Within Group Sum of Squares

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Today, the advanced technology provides the opportunity to collect vast amounts of data in the business environment. Data mining has emerged as an effective approach for the discovery of interesting and hidden patterns in the data. It combines several disciplines together including statistics, computer science, database management and machine learning. The insights gained help companies support and improve their decision making processes.

Several studies point out the importance of data mining in a business environment. A study by Columbus (2015) points out that 89% of business leaders foresee data mining as a revolution in business. Among them, 83% of them have pursued data mining projects in their organizations. Furthermore, the respondents contribute to the survey by defining one or more factors for the potential application areas of data mining in their organization. They believe that it is profitable to predict customer behaviors (46%), to predict sales (40%), and to predict fraud or financial risk (32%). Some other benefits of adopting data mining to their organizations are finding correlation in the data (48%), analysis of social network comments (29%), analysis of high-scale machine data (28%), identifying computer security risks (29%), analysis of web streams (24%), and others (1%).

IBM Research (2011) claims that, using data mining, they are successful in the detection of credit card frauds within three hours, analysis of 100 millions of PEPSICO's documents daily, analysis of the risk and stability of Wall Street hourly, filtering digital rights of 500 billion photos per year, reducing the approval time of traffic problems to two milliseconds per decision, and many others.

Another study by O'Marah et al. (2014) report a business survey which discusses the advantages of data mining in the supply chain. The report highlights 64% of respondent's interest. Also, it attracts 31% of the respondents but they are not sure about the usefulness of data mining. Only, the remaining 5% expresses a negative opinion. Some papers study the real-life applications of data mining in filtering social media (He et al. 2013), marketing (Radhakrishnan 2013), learning diseases (Austin et al. 2013), and customer relationship management (Wei et al. 2013).

Motivated by these studies, this thesis proposes a data mining based forecasting methodology for companies. Forecasting has a vital role in a company, as accurate and reliable sales forecasting is the fundamental basis for the production planning process. The adoption of data mining to forecasting innovates the traditional methods including moving average (MA), autoregressive integrated moving-average (ARIMA), exponential smoothing (ES), and Holt-Winters (HW) (Brockwell and Davis 2002). Instead of traditional time series analysis, data mining is able to recognize the hidden patterns in a dataset by measuring the similarities (Berndt and Clifford 1994, Keogh and Pazzani 2000, Chen et al. 2012, Górecki 2014, Lines and Bagnall 2015), reducing the dimensionality (Chakrabarti et al. 2002, Barrack et al. 2015), conducting segmentation (Liao 2005, Chen and Lu 2017), and finding outliers (Loureiro et al. 2004, (Murugavel and Punithavalli 2011).

In particular, data mining based forecasting is used to deal with the vast amounts of data. It facilitates forecasting process as it can handle datasets with various characteristics such as nonlinearity, outliers, intermittency, and so on. However, decision makers also consider the trade-offs between accuracy and complexity (memory requirements) to select the best technique of forecasting. When the product variety of a company increases, it is difficult to develop forecasting methods for each product. Hence, it is important to balance high accuracy and less complexity so that decision makers can apply the techniques in their organizations, and results are interpretable.

In this thesis, the aim is to develop data mining based forecasting methodology which achieves high accuracy with less complexity. In practice, the proposed methodology can be applicable to a wide variety of companies including retailers, fast fashion, and so on.

2. THEORETICAL FUNDAMENTALS AND LITERATURE REVIEW

Data mining based forecasting has been studied widely. This chapter provides several studies about estimating future trends. It is organized into four subsections. Section 2.1 discusses the importance of using an appropriate method in forecasting so that companies maintain their competitive advantage. Section 2.2 provides data mining applications in forecasting. Section 2.3 discusses the benefits of data mining based forecasting for inventory management. Section 2.4 emphasizes the major contributions of the thesis.

2.1. Forecasting

Sales forecasting is a tool used by decision makers to estimate the future outcomes based on the historical data (Mentzer and Moon 2004). This system should be designed accurately in order to improve the performance of supply chain, i.e. lower inventory cost, smoother production plans (Zhao et al. 2001), reduced stock outs (Wisner et al. 2014), satisfied customers (Moon et al. 2003), and reduced bullwhip effect (So and Zheng 2003).

There are various approaches for sales forecasting. It is important to select the appropriate method according to the data type. Choi et al. (2014) indicate that forecasting methods are selected considering their assumptions about time series data. Note that, time series data refer to the observations measured sequentially over a time horizon. For this reason, it is critical to understand the behavior of the time series (Brockwell and Davis 2002).

Some widely known methods for dealing with time series forecasting are statistical models. These techniques find the patterns of the input data in order to model a suitable equation. This category includes moving average (MA), single exponential smoothing (Brown 1959), Holt-Winters model (Winters 1960), and autoregression integrated moving average (ARIMA) (Box and Jenkins 1976). However, Boylan and Syntetos (2010) claim that the traditional methods fail in time series data with noise, outliers, intermittency, and so on.

Intermittent data is characterized as random data with a large proportion of zero values (Syntetos and Boylan 2001), and forecasting is difficult due to its high variability. Several methods are developed to handle intermittent data, such as Croston's method (Croston 1972), adjusted exponentially weighted moving average (EWMA) (Johnston and Boylan 1996), adjusted Croston's method (Syntetos and Boylan 2001), bootstrapping (Snyder 2002), modified Holt (Altay et al. 2008), and advanced Holt-Winters (Bermúdez et al. 2006).

In fact, real-life data may be non-stationary, non-linear, insufficient, and they may also include high fluctuations. To overcome these problems, data mining based forecasting methods such as support vector regression (SVR), backpropagation neural network (BPN), and cerebellar model articulation controller neural network (CMACNN) (Lu et al. 2012) have been developed.

A number of studies suggest that SVR has gained considerably wider acceptance in time series forecasting, including intermittent data (Bao et al. 2005), due to its strengths compared to other approaches (Levis and Papageorgiou 2005, Yu et al. 2013). Nalbantov et al. (2007) claim that SVR can be used to avoid overfitting problems and to improve the robustness of outlier detection. In addition, Thissen et al. (2003) explain that SVR implementation has advantages, such as finding a globally optimal solution and calculating a nonlinear solution efficiently. Das and Padhy (2012) discuss the advantage of SVR in forecasting the non-linear time series of stock market compared to the use of back propagation neural network (BPN). Zuo et al. (2014) obtain the best outcome with SVR model compared to linear discriminant analysis, logistic regression, and Bayesian network for the Radio Frequency Identification (RFID) data of consumer in-store behavior.

Hybridization of SVR with other methods improves the accuracy. Wisner et al. (2014) state that integrated forecasting is expected to reduce large errors. Hua and Zhang (2006) conclude that hybridization of logistic regression and SVR (LRSVM) outperforms the forecasting methods for intermittent time series such as Croston's method, Markov bootstrapping, and single SVR.

Some studies focus on feature selection to generate a better SVR. Lu et al. (2009) apply an independent component analysis (ICA) in order to remove the features containing noisy values. ICA together with SVR results in better accuracy in forecasting financial time series compared to pure SVR. Lu et al. (2012) also perform feature selection, and it utilizes multivariate adaptive regression splines (MARS) with SVR. In a recent study, Lu (2014) extracts additional features adopted from technical indicators of the stock market, and characterizes different properties of the data set, i.e. trend, growth, and volatility.

The details of data mining based approaches are given in Section 2.2.

2.2. Data Mining

Forecasting can become a difficult task when there is 1) no previous sales for an item (in the case of launching new items), 2) a massive sales dataset for a large number of items, and 3) a need for descriptive features to determine the customer's behavior. Thomassey (2010) claims that data mining can be used to resolve these issues.

Data mining is an effective tool for business intelligence to discover the patterns and knowledge from massive data sets (Gorunescu 2011). Sharma (2014) lists the reasons of using data mining: 1) large data with insufficient information, and 2) necessity to extract the useful information and patterns from the data.

Data mining tasks could be predictive and descriptive. Descriptive methods such as clustering and association rule mining extract the general characteristics of the dataset. Predictive methods such as classification and regression make predictions using the existing datasets.

Clustering is to partition the data set into disjoints clusters according to their similarity values (Han et al. 2012). Clustering is adopted for customer segmentation so that customers with similar characteristics and sales patterns are grouped. Therefore, some clustering algorithms have been applied for customer segmentation. Customer segmentation can be performed using 1) categorical variables, i.e. purchased frequency (Bala 2012) and customer's background (Biscarri et al. 2017), or 2) time series data (Lu and Kao 2016, Chen and Lu 2017). The algorithms used in clustering-based forecasting are hierarchical clustering (Huber et al. 2017, Biscarri et al. 2017), k-means (Kuo and Li

2016, Dai et al. 2015), fuzzy c-means (Bao et al. 2004), and association rules (Tsai et al. 2009, Xiao et al. 2011). Kuo and Li (2016) and Dai et al. (2015) apply k-means algorithm. Then, they use SVR to predict the forecasts for each cluster. Murray et al. (2017) claim that clustering task is helpful to forecast the sales of a large number of customers. Since segmenting the customers into groups based on their similar buying behaviors can simplify forecasting. Hyndman et al. (2014) support that clustering allows to handle the forecasting for large datasets due to: 1) individual prediction is too costly, and 2) aggregation of the entire models are not effective because of noise. Murray et al. (2015) emphasize that clustering customers is also convenient for examining their sales data, even when the descriptive features are not available.

In clustering, the similarities among the objects are measured using various distance functions. The Euclidean distance defined by Agrawal et al. (1993) is often used to calculate the similarity between two objects. It is used in various studies on clustering-based forecasting (Thomassey and Fiordaliso 2006, Kumar and Rathi 2011, Chen and Lu 2017). Nevertheless, Euclidean distance is not a proper function for the datasets with different lengths (Keogh 1997). For this reason, an elastic measure, dynamic time warping (DTW) (Berndt and Clifford 1994), is introduced. DTW algorithm aligns a pair of sequences by warping their vectors iteratively. It measures the cost matrix between the assigned vectors through the Euclidean distance. The goal is to achieve an optimal match, which relates the vectors in two sequences, by minimizing the total cost. There are also other measures such as regression time warping (RTW) (Lei and Govindaraju 2004), segment-wise time warping (STW) (Zhou and Wong 2005), scaled and warped matching (SWM) (Fu et al. 2008), weighted DTW (WDTW) (Jeong et al. 2011), and Shape DTW (ShapeDTW) (Zhao and Itti 2018). A comprehensive explanation of DTW can be found in Section 3.2.2.

Meanwhile, Han et al. (2012) explain that classification task has the advantage of characterizing the dataset. It identifies the data points which belong to a group. Thomassey and Fiordaliso (2006) cluster a large number of apparel items, and, then, classify them to describe the characteristics of sales data. It is helpful to determine the relations between the sales data and the descriptive criteria, which may influence the apparel sales, i.e. weather, holiday, promotions, and economic environment. In terms of prediction, C4.5 algorithm associates the new products with the closest clusters and uses its prototype to determine the future sales. Moreover, Thomassey and Happiette (2007) focus on a similar problem, and they introduce Probabilistic Neural Network (PNN) as a classifier.

Numerous studies conclude that a decision tree classifier provides benefits in analyzing customers' behaviors (Biscarri et al. 2017) and prediction (Ou and Wang 2009, Lai et al. 2009, Kirshners et al. 2010, Kumar and Rathi 2011). It could be utilized both for categorical variables (classification tree) and continuous variables (regression tree). An early algorithm for decision tree construction is ID3 (Iterative Dichotomiser) (Quinlan 1986) and followed by C4.5 (a successor of ID3) (Quinlan 1993) and Classification and Regression Tree (CART) (Breiman et al. 1984). According to Duch et al. (2004), C4.5 algorithm is widely used in many applications. However, CART algorithm is more suitable for numerical problems.

The integrated application of clustering and classification is also used in order to improve forecasting accuracy when the dataset is too large and noisy. Thomassey (2010) combines k-means clustering and decision tree to forecast sales in clothing industry. In the first task, items are segmented into clusters according to the similarity of their sales curves. It aims to reduce the complexity and noise (Witten et al. 2011). Cluster prototype, namely cluster medoid, is determined to represent the sales pattern of each cluster. In the second task, a classification model is performed for each cluster to determine the relations between the prototypes of sales and descriptive criteria. The classifier assigns a new item to one of the cluster prototypes based on its descriptive criteria. The future sales of new items are predicted through the cluster's prototype by applying an adaptive network-based fuzzy inference system (ANFIS), autoregressive moving average exogenous (ARMAX), and Holt-Winters approach.

2.3. Data Mining Based Forecasting for Inventory Management

Inventory management is the process of satisfying the customer demand on time while keeping the inventory cost at the minimum level (Coyle et al. 2003). It basically serves two goals (Reid and Sanders 2007): 1) assuring the availability of required materials, and 2) balancing customer satisfaction and total cost.

Data mining is an emerging tool for inventory management. Tsai et al. (2009) adopt agglomerative hierarchical clustering technique to learn the order demand behavior. The highly correlated items, i.e. jointly ordered, are clustered into the same group, whereas low correlated items are ordered separately. The goal is to determine the items that would be substituted for each other so that can-order policies can be applied in the joint replenishment problem. The maximum total profit is obtained from the scenarios which include clustering strategies.

Another application of data mining in inventory management is promoted by Xiao et al. (2011). They classify inventory items based on the lost profit rule. The authors develop ABC classification to distinguish the importance of items by considering not only the sales profit, but also the lost profit.

Meanwhile, Bala (2010, 2012) offers the use of data mining with forecasting to optimize the inventory level. Bala (2010) applies classification to extract the behavior of the purchased demand. Customers are segregated according to their total of purchased items. Then, their profiles are determined with a decision tree classifier, and the important factors that may affect purchasing behavior are found. Afterwards, ARIMA is used to forecast the future sales for each class. The proposed approach gives the smallest error compared to the pure ARIMA forecasting. Considering a periodic review policy, the proposed forecasting method attempts to analyze the multi-item inventory replenishment with respect to the inventory level and customer service. (Bala 2012) uses the same idea to classify the customers regarding the purchased items. The difference is that he applies classification to select the important attributes based on the target classes. He then uses the selected features, i.e. gender, income, number of children, level of education and domicile of the province, to do the clustering procedure. He considers the clustering-based forecasting to predict the sales with ARIMA method.

2.4. The Contribution of the Thesis

This thesis proposes a new framework for data mining based forecasting and inventory management. First, the items having similar sales patterns are determined using hierarchical clustering. The sales data may have unequal lengths of sequences, so different from the previous studies, we adopted the DTW as a distance measure in clustering-based forecasting. We also determined the representatives of each cluster.

Second, features are adopted from time series classification. In addition to these, new features are proposed for intermittent data. Then, feature selection is performed using MARS. Next, SVR is used for sales forecasting.

Third, the inventory performance of the proposed approach is examined in terms of total inventory cost and inventory turnover (IT).

As a summary, the contributions of this thesis are as follows:

1. A new forecasting methodology based on data mining is proposed. The proposed methodology integrates clustering, feature extraction, feature selection, and prediction tasks of data mining.
2. Different from the previous studies, we adopt the DTW as a distance measure in clustering-based forecasting.
3. New features are developed for intermittent data.

3. MATERIAL AND METHODS

This chapter explains the material and methods used in data mining-based forecasting (DMF). Section 3.1 gives information about the material studied in the thesis. Section 3.2 explains the methods used throughout the study.

3.1. Material

In this thesis, the sales dataset of a company offering a high product variety is considered. The dataset consists of several time series sequences. Each sequence denotes the amount of sales for a product, and it may have multiple zero values, called as intermittency. In the rest of the thesis, the terms dataset and time series sequence are used interchangeably.

The aim of the study is to develop a forecasting methodology with high accuracy and less complexity. High accuracy corresponds to minimum forecasting error, whereas less complexity corresponds to having less number of features (predictor variables) and forecasting models.

3.2. Methods

The methods used in this thesis are explained in the following subsections. Section 3.2.1 introduces the clustering methods. Section 3.2.2 compares the clustering performance of two dissimilarity measures. Section 3.2.3 presents multivariate adaptive regression splines to select the useful predictor variables. Section 3.2.4 explains decision trees to determine the clusters' behaviors. Section 3.2.5 exhibits support vector regression for forecasting. Section 3.2.6 presents the proposed approach. Section 3.2.7 describes the evaluation of the inventory performance.

3.2.1. Clustering

Han et al. (2012) define clustering as a task to divide the objects based on their similarities. This task includes the discovery of the hidden patterns to gain insight. Also, it simplifies the datasets by reducing the number of objects.

There are several methods for clustering such as partitioning methods, hierarchical methods, density-based methods, grid-based methods, and model-based methods. Partitioning methods directly decomposes the datasets into the given number of clusters. It starts with initial cluster centers, and uses an iterative relocation technique to move objects among groups so that partitioning improves. In general, the number of clusters is given a priori. Reversely, hierarchical methods do not require the number of clusters. Density-based methods determine clusters from the regions having higher density. Meanwhile, grid-based and model-based methods use grids and probability distributions to build clusters, respectively. In this study, hierarchical clustering is used, so it is explained in detail as follows.

Hierarchical clustering

Hierarchical clustering (HC) groups data objects into a tree of clusters. It generates a dendrogram which can be cut to a certain height to determine the desired number of clusters (Han et al. 2012). According to the hierarchical decomposition methods, there are agglomerative (bottom-up) and divisive (top-down) approaches. In the agglomerative approach, each cluster is initialized by a data object, and then clusters having closer similarity are merged until all objects are in a single cluster. The divisive version works in the opposite direction of the agglomerative version. Figure 3.1 depicts an example of dendrogram with a horizontal line which cuts the data set into four clusters.

The similarity among clusters are defined using the linkage type. Single linkage measures the minimum distance between the two objects in different clusters (Figure 3.2 (a)). Complete linkage calculates the maximum distance between the two objects in different clusters (Figure 3.2 (b)). Average linkage finds the average distances between the object pairs in different clusters (Figure 3.2 (c)).

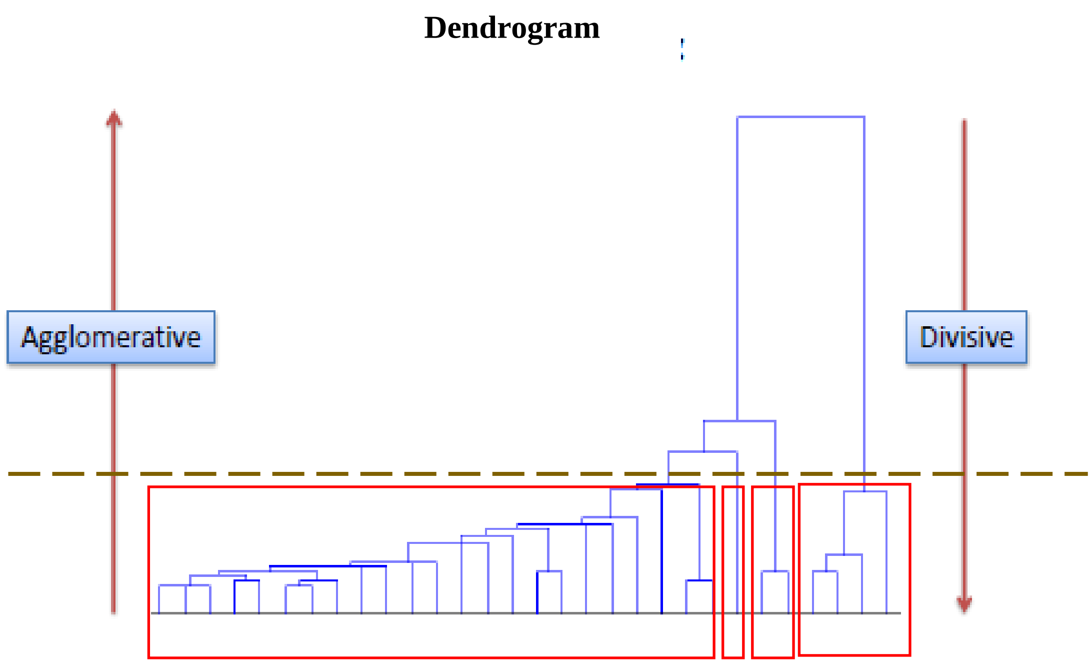


Figure 3.1. Example dendrogram (Sayad 2018)

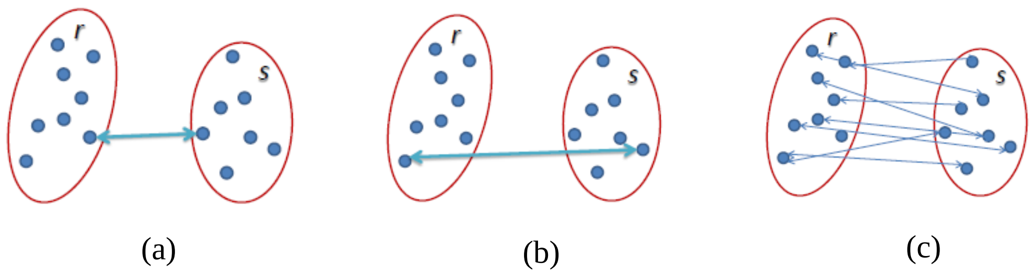


Figure 3.2. Linkage types used in hierarchical clustering, (a) single linkage, (b) complete linkage, and (c) average linkage (Sayad 2018)

HC can also be used with various distance measures including DTW. The technique to measure the distance between objects will be explained in more detail in Section 3.2.2.

Cluster prototype

A cluster prototype is the representative of the members in a cluster. Note that the term prototype is adapted from Hautamaki et al. (2008). Instead of using all members of a cluster, the cluster protoype is used to represent the characteristics of the associated cluster. In the literature, there are several approaches to obtain the cluster prototype.

Medoid approach

In time series clustering, cluster medoid is commonly used as a prototype (Hautamaki et al. 2008). That is, the data object having the minimum total distance to the other cluster members is selected as the prototype:

$$H_i = \arg\min_{S_j \in C_i} \sum_{S_k \in C_i \setminus S_j} d(S_k, S_j) \tag{3.1}$$

where H_i is the prototype for cluster i, d is the distance measure, S_k is the data object k, and C_i is the set of data objects in cluster i.

DTW barycenter averaging (DBA) approach

Another method for finding cluster prototype is DTW Barycenter Averaging (DBA) (Petitjean et al. 2011). DBA outperforms most of the existing methods of averaging, i.e. nonlinear alignment and averaging filter (NLAAF), prioritized shape averaging (PSA) (Anh and Thanh 2015), cross-words reference template (CWRT) (Soheily-Khah et al. 2015).

This approach minimizes the sum of squared DTW distances from the average sequence, namely barycenter, to the other time series sequences in the cluster. Technically, let $\mathbb{S}=\{S_1,..,S_N\}$ be the sequences of time series in the cluster, and $C = \langle C_1, C_2, \dots, C_T \rangle$ be the average sequence of $\mathbb{S}$ at iteration i. DBA starts with the initial average sequence, and it iterates so that the within group sum of squares (WGSS) with respect to the other sequences is minimized as follows:

$$WGSS(C) = \sum_{k=1}^{N} d_{DTW}^{2}(C, S_k) \tag{3.2}$$

where d_{DTW} is the DTW distance between the average sequence (C) and k^{th} sequence of time series in the cluster (S_k), and N is the number of time series sequences in the cluster.

In each iteration, two steps are performed: 1) DTW distance between the average sequence (barycenter) and each time series sequence in the cluster is computed, and 2) each coordinate in the average sequence is updated as the barycenter of the coordinates associated to it. Figure 3.3 shows four iterations of DBA on an example with two sequences.

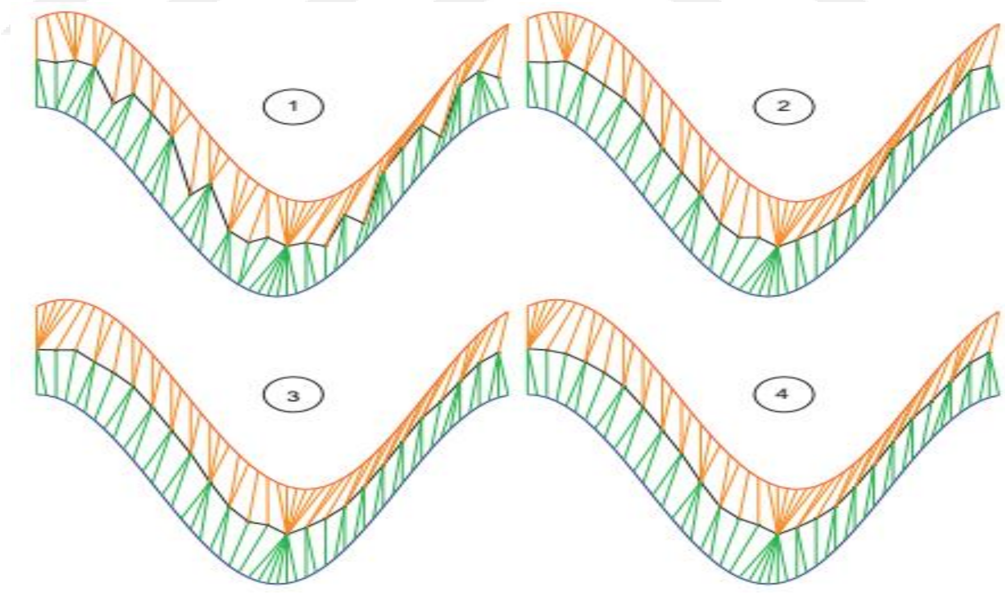


Figure 3.3. DBA iteratively adjusting the average of two sequences (Petitjean et al. 2011)

Let $C' = \langle C'_1, C'_2, \ldots, C'_T \rangle$ be the update of C at iteration (i+1). Each coordinate of the barycenter is defined in an arbitrary vector space E, $\forall \mathrm{t} \in [1, \mathrm{T}], \mathrm{C_t} \in \mathrm{E}$. The t^{th} coordinate of barycenter is then written as:

$$C'_t = barycenter(assoc(C_t)) \tag{3.3}$$

where function *assoc* links each coordinate of the average sequence to one or more coordinates of the sequences of $\mathbb{S}$, and function *barycenter* is defined as:

$$barycenter\{X_1, \ldots, X_\alpha\} = \frac{X_1 + \cdots + X_\alpha}{\alpha} \tag{3.4}$$

where X_i denotes associated coordinates and α denotes the total number of associations.

3.2.2. Dissimilarity measure

Dissimilarity measure calculates the distance between two objects. The small value of the measure indicates that the two objects have close similarity, and they can be grouped together in the same cluster. Contrarily, the high value of the measure shows the dissimilarity between two objects, so they should be assigned to the different clusters. There are several measures to define the dissimilarity among the objects.

Euclidean distance

The Euclidean distance is defined as follows (Agrawal et al. 1993):

$$d(X_i, X_j) = \sqrt{\sum_{k=1}^{n} (X_{ik} - X_{jk})^2} \tag{3.5}$$

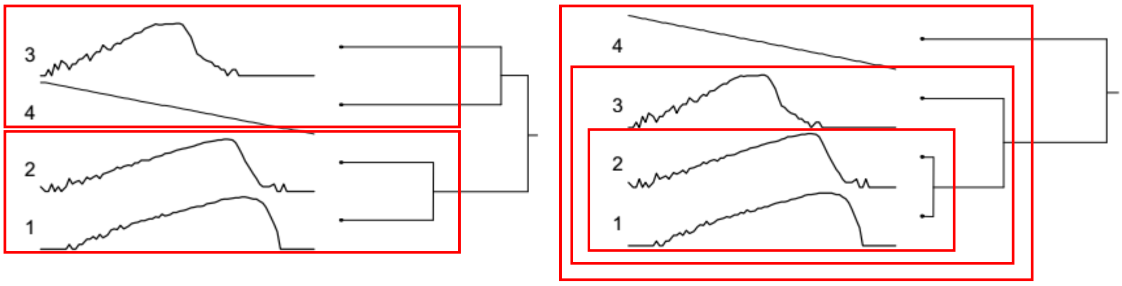
where d is the Euclidean distance between pairs, X_i and X_j are the sequences i and j, respectively, X_{ik} and X_{jk} are the kth observations of sequences i and j, and n is the length of sequence.

Euclidean distance is used in several fields such as bioinformatics (Tsai and Yu 2016), pattern recognition (Greche et al. 2017), and so on. However, it is inconvenient to use Euclidean distance under certain conditions. For example, Keogh and Pazzani (2000) show that Euclidean distance is sensitive to noise, i.e. small distortions in the time axis. Also, it calculates the similarity between a pair of sequences with equal lengths, whereas time series data may have different lengths.

Dynamic time warping

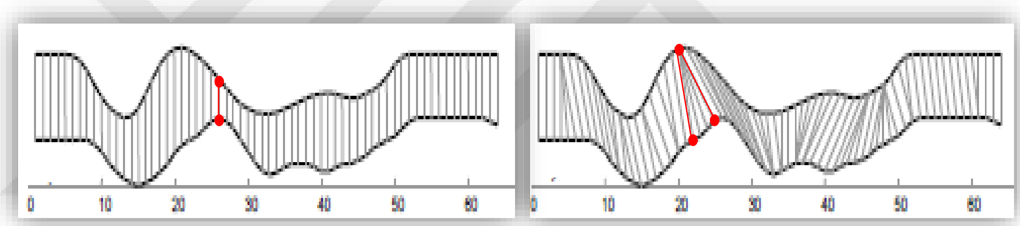
Dynamic Time Warping (DTW) calculates the dissimilarity between two sequences with unequal lengths (Berndt and Clifford 1994). Figure 3.4 (a) and (b) show the comparison of clustering results using the Euclidean and DTW distances, respectively. In Figure 3.4 (a), sequences 1 to 3 have approximately the same shape, and sequence 4 is a stright line. However, sequences 3 and 4 are considered similar using Euclidean distance. Meanwhile, Figure 3.4 (b) defines that the similarity between sequences 1 and 2 is high using DTW, and, also, these sequences are closer to sequence 3.

DTW has a sophisticated calculation to measure the distances compared to the Euclidean distance. Euclidean distance aligns i^{th} point in one sequence with i^{th} point in the other sequence (one-to-one point) (Figure 3.5 (a)). DTW extracts a warping path to align the nonlinear sequences (many-to-one or one-to many point) (Figure 3.5 (b)).



(a) (b)

Figure 3.4. Clustering results using (a) Euclidean distance and (b) DTW distance (Keogh and Pazzani 2000)



(a) (b)

Figure 3.5. Alignment between two sequences produced by (a) Euclidean distance and (b) DTW distance (Keogh and Pazzani 2000)

A warping path, W, depicts a mapping between two sequences $Q=(q_1,..,q_m)$ and $P=(p_1,..,p_n)$ of lengths m and n, respectively. Figure 3.6 illustrates the warping path W for sequences Q and P, and the matrix element (i,j) aligns q_i and p_j. Then, the k^{th} element of W is defined as $w_k=(i, j)_k$ and the warping path becomes:

$$W = w_1, \dots, w_K \qquad \max(m,n) \leq K < m + n - 1 \tag{3.6}$$

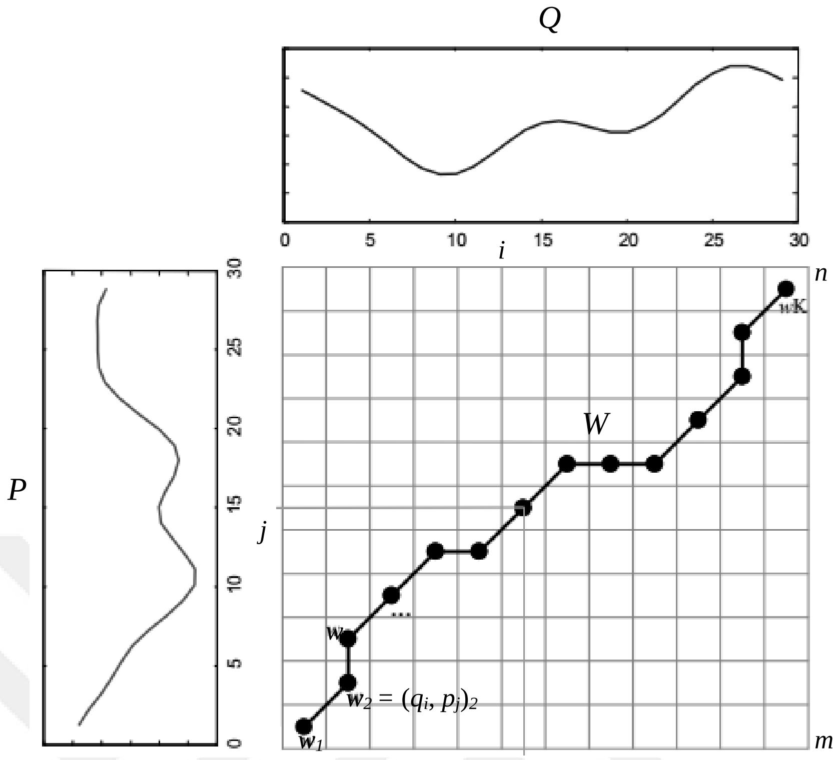


Figure 3.6. Warping path (Keogh and Pazzani 2000)

The warping path until the k^{th} element of W can be found using dynamic programming to assess the following recurrence function:

- Boundary conditions require the path to start from $w_1 = (1,1)$ and to finish at $w_k = (m, n)$ in diagonally opposite corner of matrix.

- For continuity, the allowable steps are restricted, i.e. given $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where a-$a' \leq 1$ and b-$b' \leq 1$.

- Given $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where $a - a' \geq 0$ and $b - b' \geq 0$, the points in W are forced to be monotonical.

The warping path until k^{th} element of W can be found using dynamic programming to assess the following recurrence function:

$$\gamma(i,j) = d\left(q_i, p_j\right) + \min\begin{cases} \gamma(i-1, j-1) & \\ \gamma(i-1, j) & i > 1, j > 1 \\ \gamma(i, j-1) & \end{cases} \tag{3.7}$$

where $\gamma(i,j)$ is the cumulative distance, and $d(q_i, p_j)$ is the Euclidean distance between points q_i and p_j.

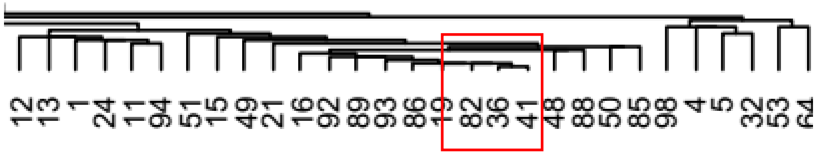
The value of warping path, W, is then minimized through a simple calculation:

$$DTW(Q,P) = \min\left\{\sum_{k=1}^{K} w_k\right\} \tag{3.8}$$

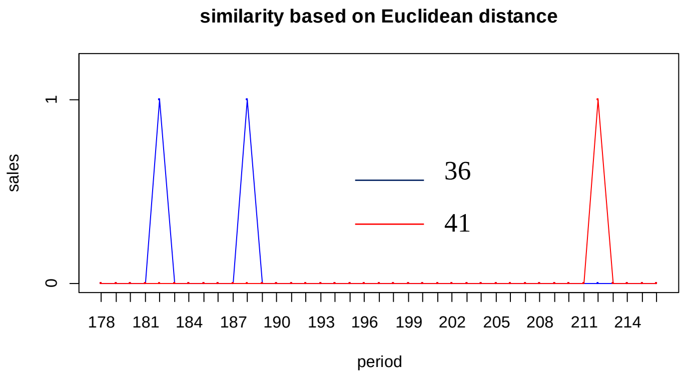
In order to calculate the distance accurately, the DTW's constraints including step pattern, window type, and window size, need to be adjusted (Giorgino 2009). Step pattern controls whether the repeated elements are consecutively matched or skipped. It can be symmetric or asymmetric. The others, i.e. window type and window size, limit warping curves to enter the certain regions of the plane. These types are illustrated as Sakoechiba (Sakoe and Chiba 1978), Itakura (Itakura 1975), and slantedband (Giorgino 2009).

Euclidean distance versus DTW

The use of Euclidean distance and DTW is compared using the dissimilarity of a pair of time series. Appendices 1 and 2 display the dendrograms for the Euclidean distance and DTW, respectively. In the dendrogram, sample items, i.e. items 82, 36, and 41, are considered. Based on the Euclidean distance, the similarity between items 36 and 41 is higher than the similarity between items 36 and 82 (Figure 3.7 (a)). However, the vice versa is true for DTW (Figure 3.8 (a)). Figure 3.7 (b) shows that the Euclidean distance does not reflect the similarity between items 36 and 41. Meanwhile, Figure 3.8 (b) denotes that items 36 and 82 have similarities. These graphs support the claim of Tormene et al. (2009) that DTW is a better similarity measure for time series data. Therefore, in this thesis, DTW is used in clustering the items.

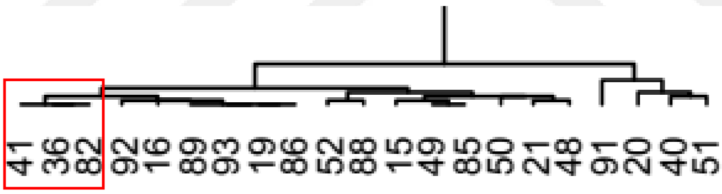


(a)

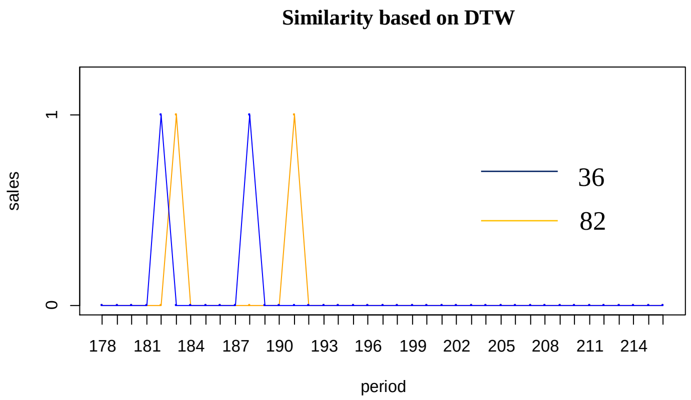


(b)

Figure 3.7. Calculation of the similarity between two items based on the Euclidean distance; (a) a sample part of the dendrogram, (b) sequences 36 and 41



(a)



(b)

Figure 3.8. Calculation of the similarity between two items based on DTW distance; (a) a sample part of the dendrogram, (b) sequences 36 and 82

3.2.3. Multivariate adaptive regression splines

Multivariate Adaptive Regression Splines (MARS) is a nonparametric regression procedure to model the interactions between dependent and independent variables without any assumption about their functional relationship (Friedman 1991). This method can handle datasets with high dimensionality. Besides, MARS can investigate the important variables without long training processes, and saves computation time (Lu et al. 2012).

MARS uses the so-called basis function (*t*-*x*) and (*x*-*t*), where *t* is the knot of the basis functions to approximate the linear or nonlinear relationships (Figure 3.9). Only positive part of the basis functions is considered, otherwise it takes a value of zero.

$$(x-t)_+ = \begin{cases} x-t & ,x>t \\ 0 & ,\text{otherwise} \end{cases} \tag{3.9}$$

where *x* is the predictor variable, and *t* is a univariate knot.

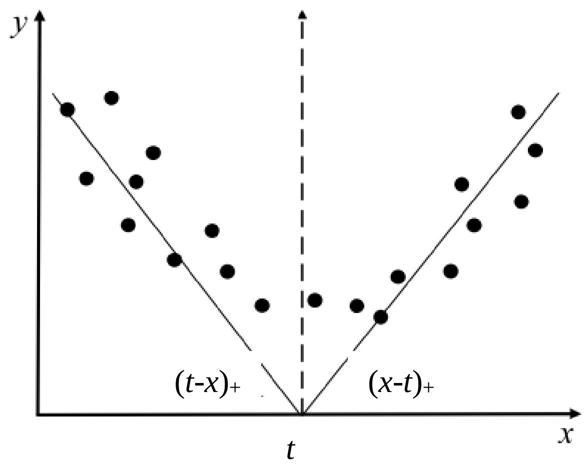


Figure 3.9. Piecewise linear basis function (Taylan and Yerlikaya-Özkurt 2010)

The general MARS function can be defined as follows:

$$\hat{f}(x) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m} [S_{km}(x(k,m) - t_{km})] \tag{3.10}$$

where a_0 is the intercept, a_m is the coefficient of the model, M is the number of basis functions, K_m is the number of knots, S_{km} is the right/left position of the associated step function, $x(k,m)$ is the label of the independent variable, and t_{km} is the knot location.

The technique starts with the simplest model of the basis function. It is followed by adding the basis function (for each variable and for all possible knots) recursively so that prediction error is minimized. This is called forward stepwise. It stops when M_{max} is reached. Then, it continues with backward procedure to fix the overfitting. It decreases the complexity without degrading the fit, and removes basis functions that contributes the smallest increase in the residual squared error. It produces an optimal estimated model $\hat{f}_\alpha$ with respect to the number of terms, α. Generalized cross validation (GCV) is used to estimate the optimal value of α as follows:

$$GCV = \frac{\sum_{i=1}^{N}(y_i - \hat{f}_\alpha(x_i))^2}{(1 - \frac{M(\alpha)}{N})^2} \tag{3.11}$$

where y_i is the response variable, x_i is the predictor variable, N is the number of sample observations in the dataset, $M(\alpha) = u+dK$ with u is the number of independent basis function, K is the number of knots selected in the forward process, and d is the penalty for adding basis function.

3.2.4. Decision trees

Decision tree is a widely used supervised learning method (Han et al. 2012). It can be used to predict both categorical and numerical class labels. It begins with a root node and grows by splitting the training set into smaller subsets according to the attribute selection measure (internal node). It ends with the leaf nodes that show the class label or function (Figure 3.10).

Commonly used measures to select the best attribute for splitting are information gain (entropy), Gini index, and classification error (Tan et al. 2006). The measures are defined as follows:

$$Entropy = -\sum_{i=1}^{C} p_i \times log_2(p_i) \tag{3.12}$$

$$Gini = 1 - \sum_{i=1}^{C} (p_i)^2 \tag{3.13}$$

$$classification\ error = 1 - \max_i(p_i) \tag{3.14}$$

where C denotes the total number of classes and p_i is probability of belonging to class i.

These measures are based on the degree of the child node's impurity. The attribute with the lowest impurity is used in the process of splitting. The splitting process is repeated until a stopping criterion is satisfied. Maimon and Rokach (2005) describe the stopping rules as follows: 1) all points belong to the same class, 2) the maximum tree depth is reached, and 3) the impurity value in a node is less than a threshold.

The algorithms such as ID3 and C4.5 use entropy to select the attributes for splitting, whereas some algorithms like CART use Gini index.

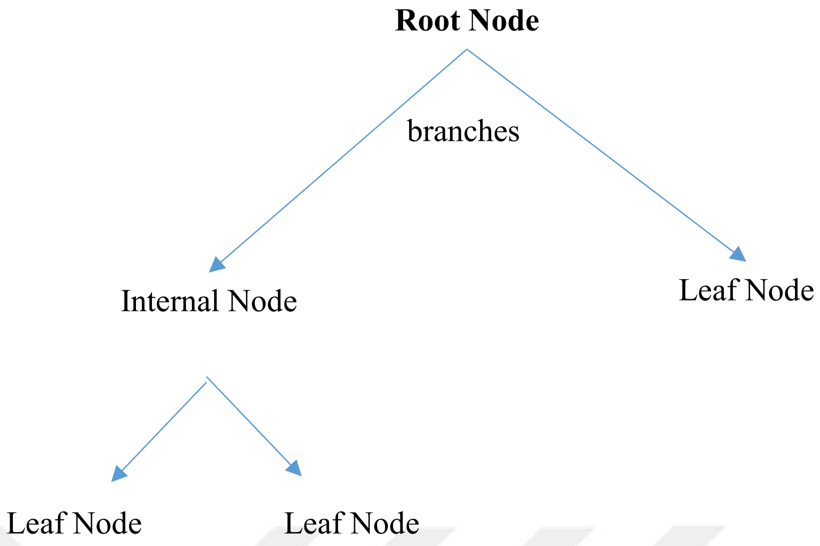


Figure 3.10. Decision tree

3.2.5. Support vector regression

Support vector regression (SVR) popularized by Vapnik (1998) uses the concept of support vector machine (SVM) to forecast the nonlinear and high dimensional problems. It is based on determining the loss function called ε-insensitivity to penalize errors.

SVR can be formulated as (Vapnik 1998):

$$f(x) = (w \cdot \emptyset(x)) + b \tag{3.15}$$

where w is a weight vector, x is the model input, $\emptyset(x)$ is a kernel function to transform the nonlinear inputs to linear form, and b is a bias.

The aim is to find a function $f(x)$ that deviates at most ε from the target values in the training data $\{(x_1, y_1),..,(x_n, y_n)\} \subset \mathbb{R}$. The slack variables ξ_i and ξ_i^* allow errors beyond ε precision. Therefore, the weight vector (w) and bias (b) are estimated using a convex optimization problem as follows:

Minimize:

$$z = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{3.16}$$

Subject to:

$$\begin{cases} y_i - \big(w \cdot \phi(x_i)\big) - b \leq \varepsilon + \xi_i \\ \big(w \cdot \phi(x_i)\big) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad \text{for } i = 1, \dots, n \end{cases} \tag{3.17}$$

where $C > 0$ is a constant coefficient to specify the trade-off between $\|w\|^2$ (flatness of function *f*) and the tolerance to deviations larger than ε.

The parameters *C* and ε-insensitivity are determined by the user. Several metaheuristics have been applied to help determining the SVR parameters, like genetic algorithm (GA) (Wu 2010), particle swarm optimization (PSO) (Safarzadegan Gilan et al. 2012), differential algorithm (DA) (Wang et al. 2012), and firefly algorithm (FA) (Xiong et al. 2014).

Using Lagrangian multipliers and Karush-Kuhn-Tucker conditions, Equations (3.16) and (3.17) transform into the dual Lagrangian form as follows (Lu 2014):

Maximize:

$$\begin{aligned} L_d(\alpha, \alpha^*) = {} & -\varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) + \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)y_i \\ & - \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)\big(\alpha_j^* - \alpha_j\big)\,K\big(x_i, x_j\big) \end{aligned} \tag{3.18}$$

Subject to*:*

$$\begin{cases} \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) = 0 \\ 0 \leq \alpha_i \leq C, \qquad i = 1, \dots, n \\ 0 \leq \alpha_i^* \leq C, \qquad i = 1, \dots, n \end{cases} \tag{3.19}$$

where α_i and α_i^* are the Lagrangian multipliers that satisfy $\alpha_i \alpha_i^* = 0$, and $K(x, x_i)$ is the Kernel function. The optimal weight vector becomes $w^* = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) K(x, x_i)$. Hence, the general function of SVR can be written as:

$$f(x, w) = f(x,\ \alpha,\ \alpha^*) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) K(x, x_i) + b \tag{3.20}$$

The commonly used kernel is the radial basis function (RBF) which is defined as:

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{3.21}$$

where σ denotes the width of the RBF.

Figure 3.11 shows an example for the transformation of the nonlinear inputs to linear form.

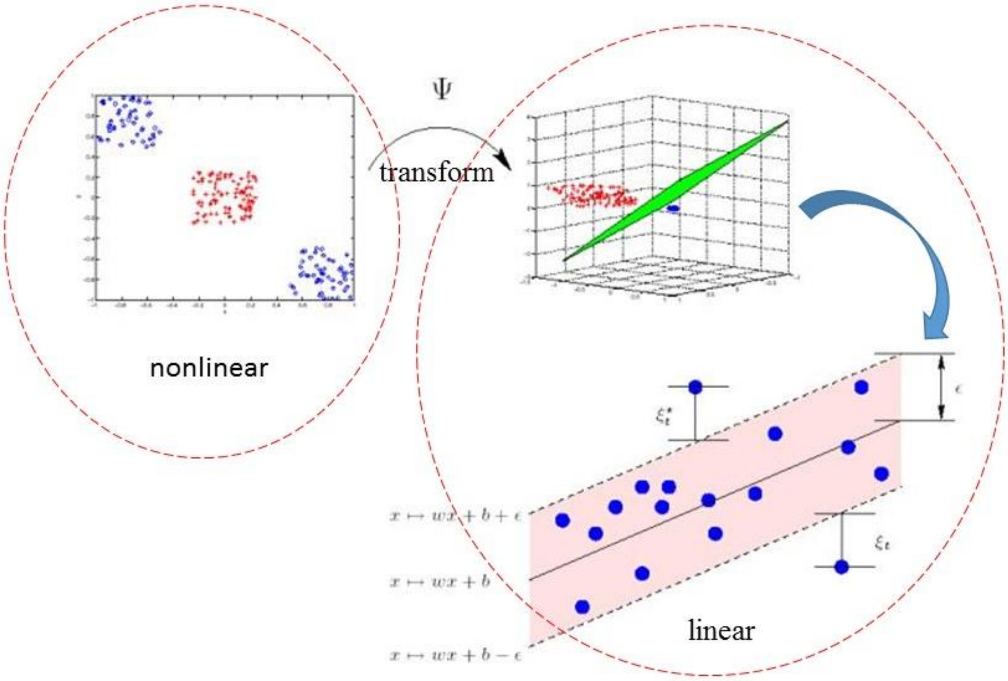


Figure 3.11. Transformation of the nonlinear problem to linear form in SVR (KernelSVM, 2018)

3.2.6. Proposed Methodology

This thesis proposes a data mining-based forecasting (DMF) methodology for time series data with unequal lengths and intermittency. It aims to achieve high forecasting accuracy using less complex models and improve the inventory performance.

The flowchart of the proposed methodology is provided in Figure 3.12. In Step 1, the sales data are collected. In Step 2, preprocessing operations are performed. That is, the inconsistencies in the data set are cleaned. The products having no sales within the planning horizon are removed. Also, the sales data of each product are cropped according to the release and phase-out times. In Step 3, products with similar sales patterns are identified using hierarchical clustering. The dissimilarities among the product sales are calculated using DTW. The number of clusters is determined using the inter-cluster heterogeneity and intra-cluster homogeneity. In each cluster, the cluster prototype is found by calculating cluster's medoid and DBA. In Step 4, the features are extracted for forecasting. In addition to the features proposed by Lu (2014), four new features are introduced. Table 3.1 lists the features proposed by Lu (2014). They characterize the amount, trend, growth, and volatility. Table 3.2 lists the proposed features to identify the intermittency. IML is calculated as the ratio of the number of zero values to the number of periods, and it considers the long-term intermittency. In IMM, first, subsequences are formed in the time series such that a positive value precedes zero value(s) in the subsequence, and the subsequence ends with a positive value. Then, the moving average of their lengths are calculated. Since the last two subsequences are considered for the moving average, IMM defines the mid-term intermittency. In order to define the short term intermittency, IMS1 and IMS2 are proposed. IMS1 calculates the ratio of the recent subsequence length to the number of zero values in the recent subsequence. Different from IMM and IMS1, IMS2 starts a subsequence with a zero value (right after a positive value). In IMS2, the number of zero values in the recent subsequence is divided by the recent subsequence length. Basically, these features show the cyclic structure of the zero demand and positive demand in the short, mid and long terms. In Step 5, MARS is used to select the important features. In Step 6, characteristics of the clusters are specified using decision tree. In Step 7, SVR is used to build a forecasting model for each cluster's prototype. The best forecasting method is selected according the accuracy and complexity. In the last step, the performance of the proposed method is evaluated in terms of inventory performance measures.

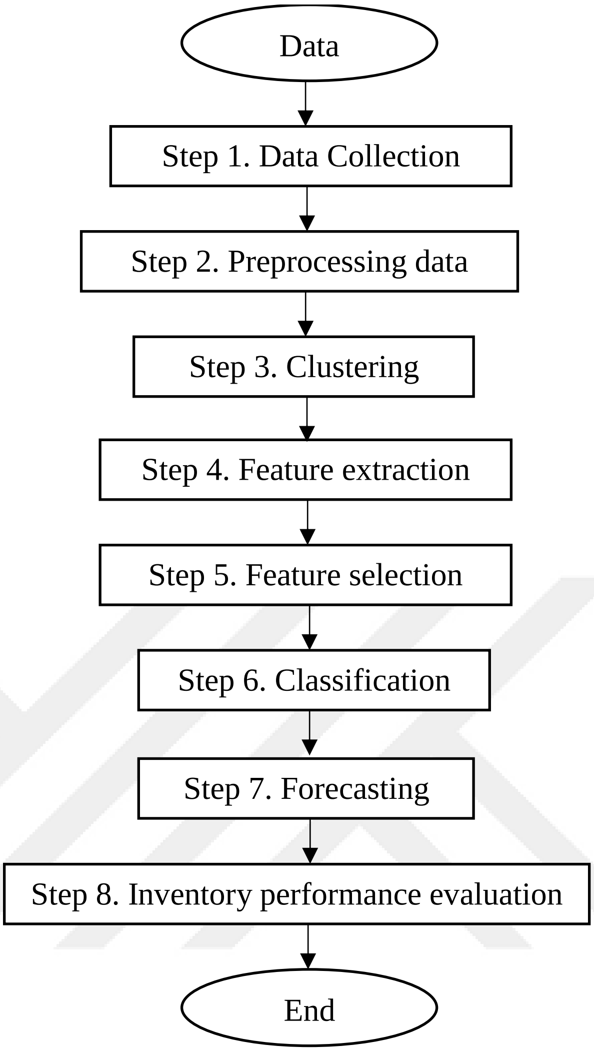


Figure 3.12. Flowchart of the proposed methodology

Table 3.1. List of the features (Lu 2014)

Variable	Description	Period	Characteristic
T1	$X_1 = C_{(t-1)}$	Short term	Amount
T2	$X_2 = C_{(t-2)}$	Short term	Amount
T3	$X_3 = C_{(t-3)}$	Short term	Amount
T5	$X_4 = C_{(t-5)}$	Mid term	Amount
T10	$X_5 = C_{(t-10)}$	Mid term	Amount
T15	$X_6 = C_{(t-15)}$	Long term	Amount
T20	$X_7 = C_{(t-20)}$	Long term	Amount
MA2	$X_8 = \sum_{i=1}^{2} C_{(t-i)}/2$	Short term	Trend
MA3	$X_9 = \sum_{i=1}^{3} C_{(t-i)}/3$	Short term	Trend
MA5	$X_{10} = \sum_{i=1}^{5} C_{(t-i)}/5$	Mid term	Trend
MA10	$X_{11} = \sum_{i=1}^{10} C_{(t-i)}/10$	Mid term	Trend
MA15	$X_{12} = \sum_{i=1}^{15} C_{(t-i)}/15$	Long term	Trend
RDP1	$X_{13} = \frac{C-C_{(t-1)}}{C_{(t-1)}} \times 100$	Short term	Growth ratios
RDP3	$X_{14} = \frac{C_t-C_{(t-3)}}{C_{(t-3)}} \times 100$	Short term	Growth ratios
RDP5	$X_{15} = \frac{C_t-C_{(t-5)}}{C_{(t-5)}} \times 100$	Mid term	Growth ratios
RDP10	$X_{16} = \frac{C_t-C_{(t-10)}}{C_{(t-10)}} \times 100$	Mid term	Growth ratios
RDP15	$X_{17} = \frac{C_t-C_{(t-15)}}{C_{(t-15)}} \times 100$	Long term	Growth ratios
BIAS5	$X_{18} = \frac{C_t-MA5}{MA5}$	Mid term	Volatility
BIAS10	$X_{19} = \frac{C_t-MA10}{MA10}$	Mid term	Volatility
BIAS15	$X_{20} = \frac{C_t-MA15}{MA15}$	Long term	Volatility
ROC5	$X_{21} = \frac{C_t}{C_{(t-5)}} \times 100$	Mid term	Volatility
ROC10	$X_{22} = \frac{C_t}{C_{(t-10)}} \times 100$	Mid term	Volatility
ROC15	$X_{23} = \frac{C_t}{C_{(t-15)}} \times 100$	Long term	Volatility
Disparity5	$X_{24} = \frac{C_t}{MA5} \times 100$	Mid term	Volatility
Disparity10	$X_{25} = \frac{C_t}{MA10} \times 100$	Mid term	Volatility
OSCP5	$X_{26} = \frac{MA5-MA10}{MA5} \times 100$	Mid term	Volatility

C_t denotes the amount of sales in period t.

Table 3.2. List of the proposed intermittency features

Variable	Description	Period	Characteristic
IML	$X_{27} = \frac{\sum_{k=1}^{t-1} I_k}{(t-1)}$	Long term	Intermittency
IMM	$X_{28} = \frac{\left(\lvert CC_{(t-1)}\rvert + \lvert CC_{(t-2)}\rvert\right)}{2}$	Mid term	Intermittency
IMS1	$X_{29} = \frac{\lvert CC_{(t-1)}\rvert}{\sum_{k\in\{k':C_{k'}\in CC_{(t-1)}\}} I_k}$	Short term	Intermittency
IMS2	$X_{30} = \frac{\sum_{k\in\{k':C_{k'}\in CP_{(t-1)}\}} I_k}{\lvert CP_{(t-1)}\rvert}$	Short term	Intermittency

Notes:
1. I_k is an indicator variable such that $I_k = 1$ if $C_k = 0$, i.e. the amount sales in period *k* is 0.
2. CC_t is a sequence $CC_t = (C_{t-k}, \dots, C_t)$ such that $C_{t-k} > 0$ and $\sum_{k'=0}^{k+1} C_{t-k'} = 0$
3. CP_t is a sequence $CP_t = (C'_{t-k}, \dots, C'_t)$ such that $C'_{t-k} > 0$ and $\sum_{k'=1}^{k} C'_{t-k'} = 0$

3.2.7. Evaluation of the Inventory Performance

The proposed approach is evaluated in terms of the inventory performance. The aim is to determine the inventory management policies that couple with the proposed forecasting approach. In this context, safety stock levels and lot sizing strategies are evaluated.

The flowchart of the inventory management procedure is presented in Figure 3.13. Note that, the proposed forecasting approach should be provided as an input to the procedure. The procedure is iterated in a given planning horizon. The initial inventory level and backordered units in the beginning of the planning horizon are provided. In the beginning of each period, the inventory levels are updated using the ending inventory level of the previous period and the received order in the current period:

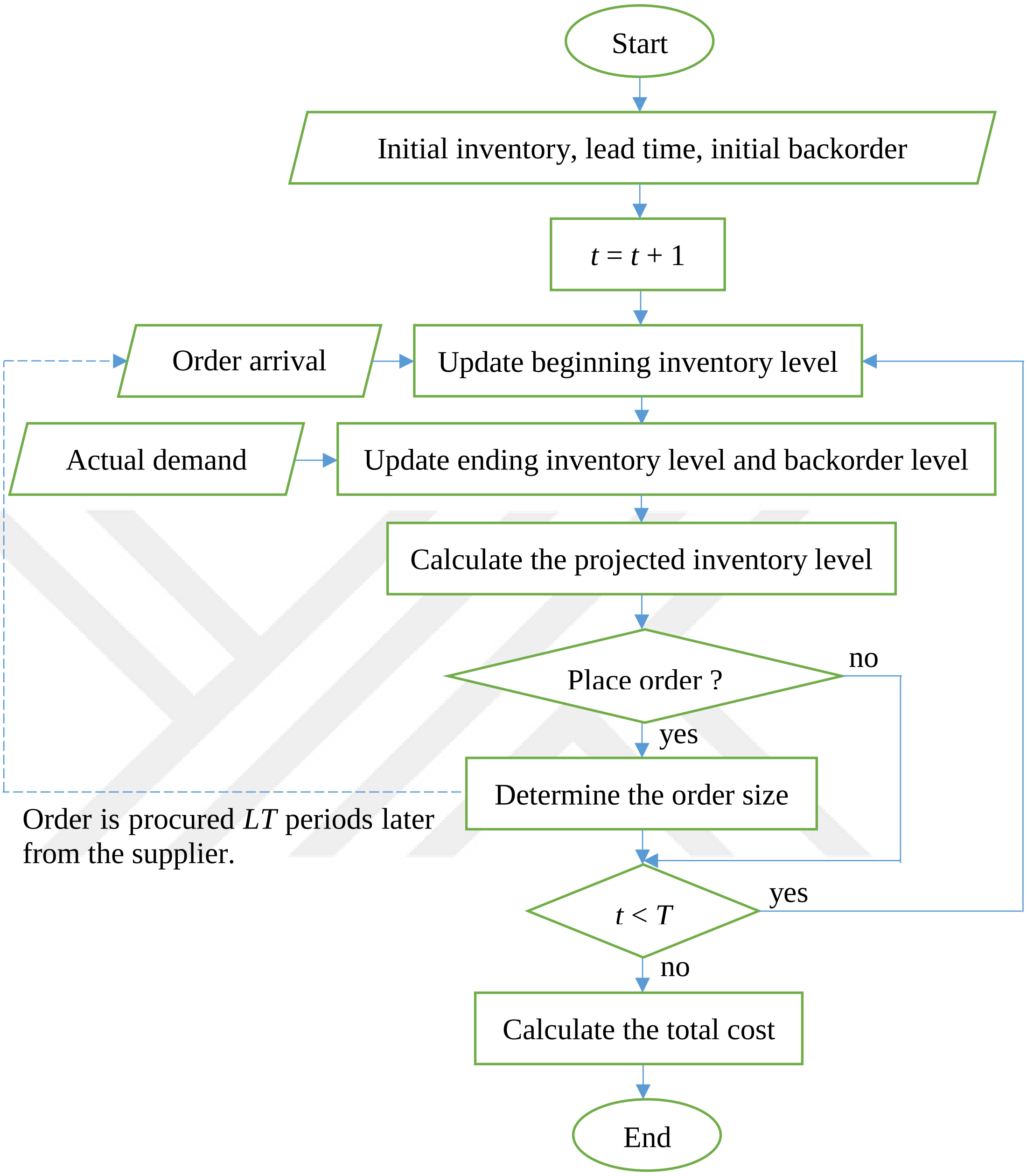


Figure 3.13. Flowchart of the inventory management procedure

$$beg.inv(t) = \begin{cases} init.inv + order(t - LT - 1) & ,\text{if } t = 1 \\ end.inv\ (t-1) + order(t - LT - 1) & ,\text{if } t > 1 \end{cases} \tag{3.22}$$

where *beg.inv*(*t*) is the inventory level in the beginning of period *t*, *init.inv* is the initial inventory level in the beginning of the planning horizon, *end.inv*(*t*-1) is the inventory level in the end of period (*t*-1), LT is the lead time, and *order*(*t*-LT-1) is the order placed at the end of period (*t*-LT-1) and received LT periods later (at the beginning of period *t*).

During period *t*, actual demand is realized, and it is satisfied from the inventory or the customers backorder for the next periods. *B*(*t*) denotes the backordered units in the end of period *t*, and *D*(*t*) is the actual demand in period (*t*). *B*(*t*) is updated as follows:

$$B(t) = \begin{cases} 0 & ,\text{if } beg.inv(t) > D(t) + B(t-1) \\ D(t) + B(t-1) - beg.inv(t) & ,\text{if } beg.inv(t) \leq D(t) + B(t-1) \end{cases} \tag{3.23}$$

At the end of the period, the ending inventory level is calculated as follows:

$$end.inv(t) = \begin{cases} beg.inv(t) - \big(D(t) + B(t)\big) & ,\text{if } B(t) = 0 \\ 0 & ,\text{otherwise} \end{cases} \tag{3.24}$$

Projected inventory position (IP) at the beginning of period (t+LT+1) ($ProjectedIP(t + LT + 1)$) is evaluated using the ending inventory, orders that will be received during the lead time, backordered units and demand during the lead time. Since actual demand is not known during the lead time, forecasts are performed using the proposed forecasting methodology.

$$ProjectedIP(t + LT + 1) = end.inv(t) + \sum_{t'=t+1}^{t+LT} order(t') - B(t) - \sum_{t'=t+1}^{t+LT+1} FD(t') \tag{3.25}$$

where $FD(t)$ is the forecast in period t. Given a safety stock level of SS, order placement decision is made as follows:

$$PO(t) = \begin{cases} 1 & , \text{if } ProjectedIP(t + LT + 1) < SS \\ 0 & , \text{otherwise} \end{cases} \tag{3.26}$$

If an order is placed ($PO(t)$=1), the order size is calculated using the forecasts beyond the lead time. For lot-for-lot and fixed w-period methods, the order sizes are determined as follows:

$$order(t) = \begin{cases} FD(t + LT + 1) & , \text{if lot} - \text{for} - \text{lot (LFL) is used} \\ \sum_{t'=t+LT+1}^{t+LT+w} FD(t') & , \text{if fixed w} - \text{period is used} \end{cases} \tag{3.27}$$

In the lot-for-lot method, the order size is equal to the forecast in the corresponding period, whereas forecasts for the next w periods are combined in the fixed w-period method.

The inventory performance are evaluated with respect to the total cost and the ratio of inventory turnover (IT). The total cost includes the ordering cost, holding cost, and backordering cost. IT denotes the ratio of the total sales in the planning horizon to the average inventory level. The aim is to minimize total cost and maximize IT.

$$TC = \left(K \times \sum_{t=1}^{T} PO(t) \right) + \frac{h}{2} \sum_{t=1}^{T} (beg.inv(t) + end.inv(t)) + b \sum_{t=1}^{T} B(t) \tag{3.28}$$

$$IT = \frac{\sum_{t=1}^{T} D(t)}{\sum_{t=1}^{T} (beg.inv(t) + end.inv(t))/2T} \tag{3.29}$$

where K is the ordering cost, h is the holding cost per period per item, and b *is* the backordering cost per period per item.

4. RESULTS

This section presents the implementation of the proposed approach in a real-life problem. It includes the company overview (Section 4.1), the sales data (Section 4.2), the parameters and performance measures (Section 4.3), and numerical results (Section 4.4).

4.1. Company's Overview

In this thesis, PT Traktor Nusantara, a distributor company of industrial heavy equipment, i.e. forklifts, is considered. The company is located in Indonesia, and it has the biggest market share for the forklift distribution, 40%, for the last five years with 15 branch offices throughout the country. It employs 750 workers to satisfy the demand of about 5,000 customers.

4.2. Sales Data

The data set used in this thesis consists of 100 time series sequences, and each sequence denotes the monthly sales amount of a forklift. Sales data span a time horizon of 19 years, from January 1998 to December 2016. Figure 4.1 shows the sales data of four example products.

These sequences may have different lengths and patterns as shown in Figure 4.2, item 60-8FD25 had a bell-shaped sales pattern, and it was sold on the market between 2010 and 2016. Meanwhile, item 60-8FD15 had a flat sales pattern and it was sold on the market between 2011 until 2016.

The intermittency of each item is evaluated. The intermittency level for item i, $ilevel(i)$, is calculated as follows:

$$ilevel(i) = \frac{\text{\# of zero demand values in item } i}{\text{length of sales data for item } i} \tag{4.1}$$

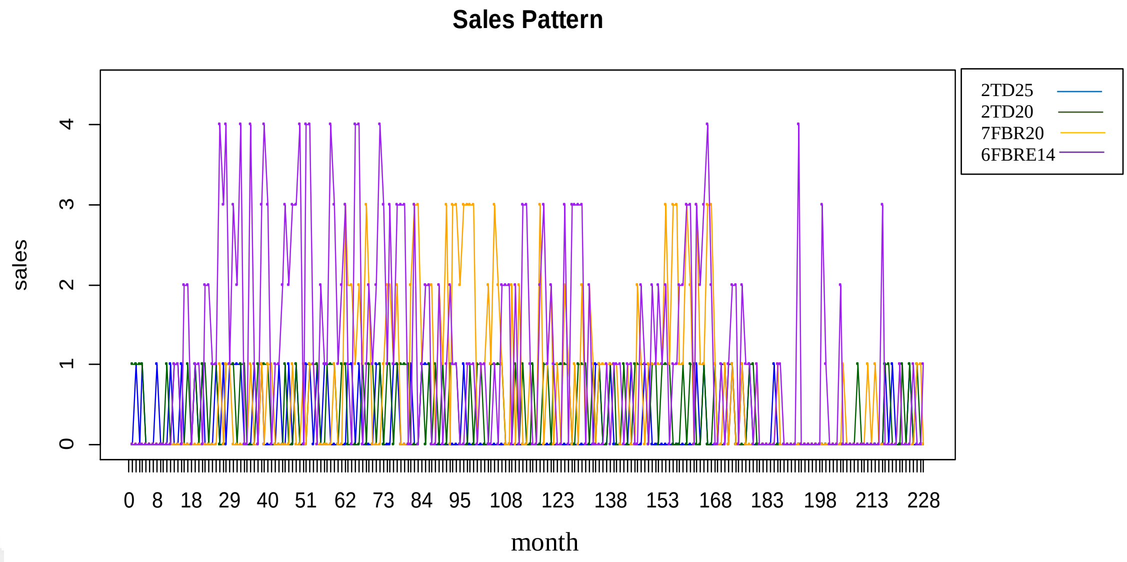


Figure 4.1. Sales pattern of four example products

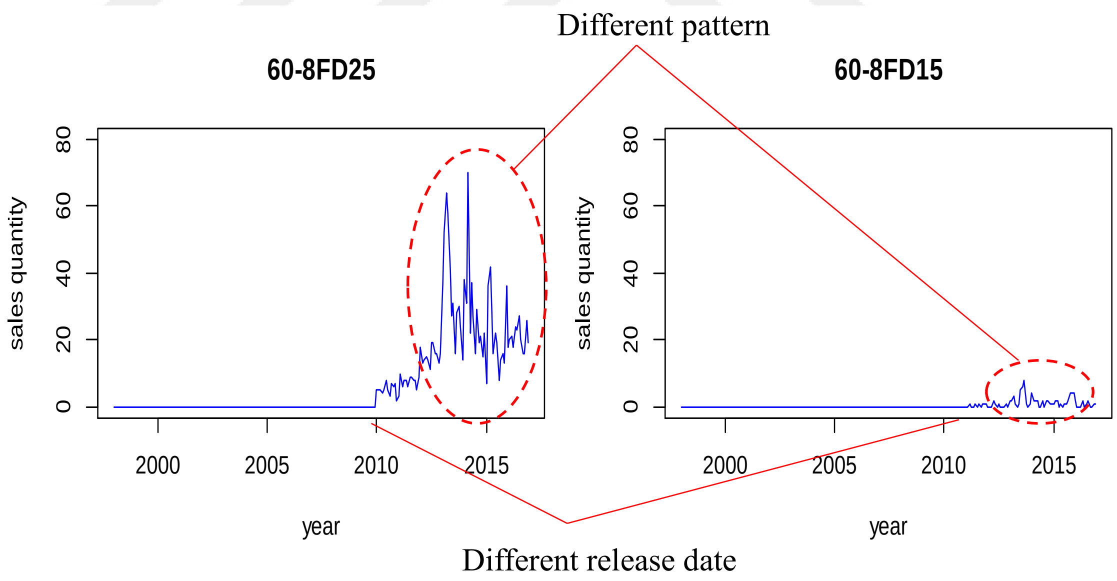


Figure 4.2. Times series sequences for items 60-8FD25 and 60-8FD15

Figure 4.3 presents the histogram of the intermittency levels for all products. The histogram indicates that more than half of the products show high intermittency, i.e. the intermittency level is higher than 0.5. Only two products have no intermittency.

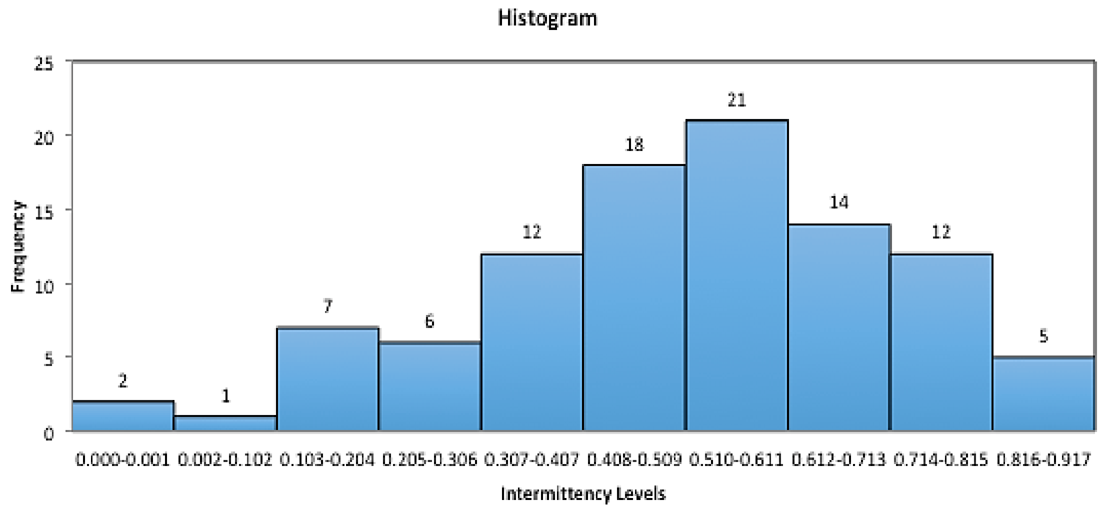


Figure 4.3. Histogram of the intermittency levels for all products

4.3. Parameter Settings and Performance Criteria

In the proposed methodology, agglomerative hierarchical clustering with complete linkage is used. The parameters of DTW are set as follows: i) step pattern is symmetric2, ii) window type is slantedband, and iii) window size is 16.

The performances of the decision tree, MARS, and SVR are evaluated using leave-one-out cross-validation. In the decision tree, no pruning method is used.

In SVR, RBF kernel function is selected according to e1071 package (Meyer et al. 2017). Grid search is applied to determine the best parameter settings for C and ε. In the search, C is changed within a range of $[2^0, 2^{15}]$ with step size 2, and ε is changed within a range of [0, 1] with step size 0.01.

The forecasting results are evaluated using mean absolute percentage error (MAPE), root mean square error (RMSE), mean square error (MSE) and mean absolute deviation (MAD) (Lu 2014). The small values of error performance measures indicate that the predictied values are closer to the actual values. The formulas of the performance measures are as follows:

$$MAPE = 100 \times \frac{1}{n}\sum_{k=1}^{n}\frac{|y_k - \hat{y}_k|}{y_k} \quad (4.2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_k - \hat{y}_k)^2} \quad (4.3)$$

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(y_k - \hat{y}_k)^2 \quad (4.4)$$

$$MAD = \frac{\sum_{k=1}^{n}|y_k - \hat{y}_k|}{n} \quad (4.5)$$

where n is the number of periods, y_k is the actual demand for the k^{th} period, and $\hat{y}_k$ is the predicted demand for the k^{th} period.

4.4. Numerical Results

The numerical results of the proposed methodology are explained in this section. All the analysis in this thesis is conducted using the R statistical computing software (R Core Team 2017) and Microsoft Excel software.

4.4.1. Preprocessing results

The time series sequences having all zero values in the planning horizon were removed. Hence, two types of forklifts with no demand are excluded, and 98 sequences are used in the analysis. Also, the time series sequences were cropped according to the release and phase-out times of each item. The resulting time series have lengths varying between 12 and 228.

Each value in the time series denotes the amount of sales, so normalization is not applied to the data set.

4.4.2. Clustering results

Determination of the number of clusters is a challenging task, as there is not a widely accepted method in the literature (Jain 2010). Since the aim of this study is to obtain homogenous clusters with products having similar sales pattern, clustering results are evaluated according to both homogeneity and heterogeneity measures. Homogeneity is calculated as the mean pairwise DTW distance within the same cluster, whereas heterogeneity is calculated as the mean pairwise DTW distance between two clusters. While the value of heterogeneity is a "larger-the-better" measure, the value of homogeneity is regarded as a "smaller-the-better" measure.

In this thesis, the number of clusters (k) is varied between 2 and 30, and agglomerative clustering algorithm is applied with DTW distance. The heterogeneity and homogeneity measures are plotted with respect to the number of clusters in Figure 4.4. The result indicates that homogeneity and heterogeneity measures stabilize for k = 7, 16, 27.

The dendrogram for k=7 is presented in Figure 4.5. In the clustering result, there are four singletons, i.e. four clusters with a single item. Also, clusters 1, 2 and 6 have 70, 21 and 3 members, respectively. For k=7, k= 16, and k= 27, the cluster assignments of the items are shown in Appendix 3.

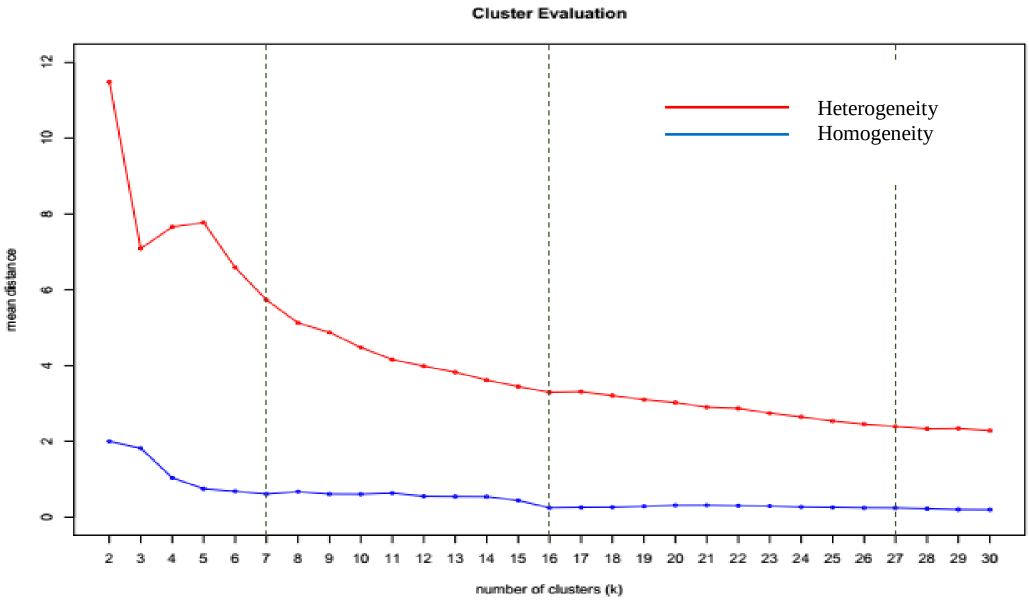


Figure 4.4. Evaluation of the number of clusters with respect to homogeneity and heterogeneity measures

Figure 4.6 shows the sequences assigned to each cluster and the prototype of each cluster for *k*=7 and DBA. In Figure 4.6, the cluster prototypes represent the average of the patterns in the clusters. For the singleton clusters, the single cluster member is assigned as the prototype.

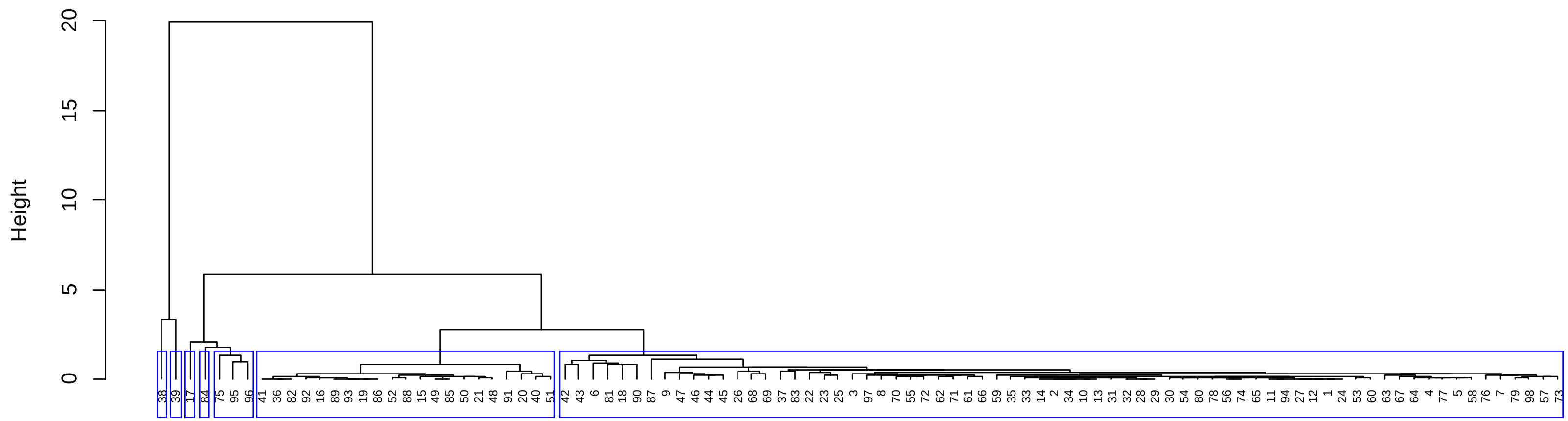


Figure 4.5. Dendrogram for k=7 (blue rectangles show the seven clusters)

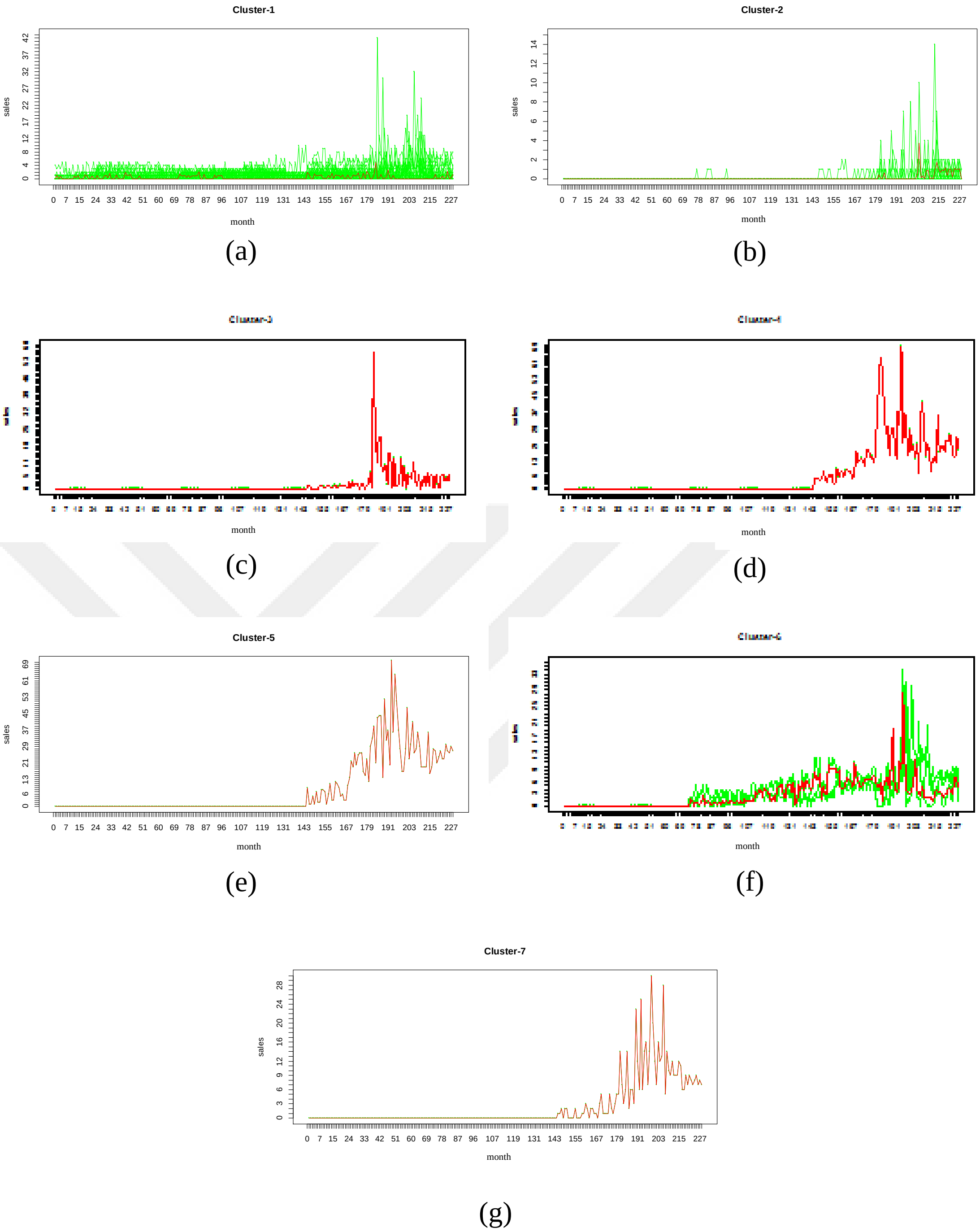


Figure 4.6. Cluster members for *k*=7 and DBA setting (red lines show the cluster representatives)

4.4.3. Feature extraction and selection results

The forecasting performances with and without intermittency features are compared in order to evaluate the contribution of the proposed intermittency features in Table 3.2. In the forecasting without intermittency features (IF), only features in Table 3.1 are used as input to the SVR model, whereas, in the forecasting with IF, features in both Tables 3.1 and 3.2 are used in the SVR.

The summary of the results of individual forecasting with and without intermittency features is presented in Table 4.1. For the products with positive intermittency level (96 products), it indicates that SVR with IF reduces the average error performances except MSE. Also, the maximum value and the standard deviation of the all error performances are increasing. It is due to the existence of an outlier, i.e. product with an extreme fluctuation. When the outlier is removed, it is observed that MAPE, RMSE, MSE, and MAD values for SVR with IF are less than the ones without IF. Hence, the proposed features are able to characterize the intermittency of time series, and improve the forecasting errors.

After feature extraction, MARS from earth package (Milborrow 2014) is applied to select the useful features. For a sample product, 16 out of 30 features are selected: *T*10, *T*20, *MA*3, *MA*5, *MA*15, *RDP*3, *BIAS*10, *BIAS*15, *ROC*10, *Disparity*5, *Disparity*10, *OSCP*5, *IML*, *IMM*, *IMS*1, *IMS*2.

The selected features for the prototypes of 7-, 16- and 27-cluster are provided in Appendices 4 and 5.

Table 4.1. Evaluation of the proposed intermittency features

Performance criteria		Products with positive intermittency level			
		with outlier		without outlier	
		SVR without IF	SVR with IF	SVR without IF	SVR with IF
MAPE (%)	**Maximum**	83.047	94.650	83.047	75.769
	Minimum	0.000	0.000	0.000	0.000
	Average	27.990	19.553	28.178	18.762
	Standard deviation	21.721	22.013	21.757	20.729
RMSE	**Maximum**	6.436	10.481	4.355	4.121
	Minimum	0.000	0.000	0.000	0.000
	Average	0.969	0.810	0.911	0.709
	Standard deviation	0.916	1.176	0.728	0.622
MSE	**Maximum**	41.422	109.857	18.969	16.979
	Minimum	0.000	0.000	0.000	0.000
	Average	1.778	2.024	1.361	0.889
	Standard deviation	4.868	11.312	2.689	2.070
MAD	**Maximum**	2.656	4.810	2.656	1.923
	Minimum	0.000	0.000	0.000	0.000
	Average	0.547	0.413	0.541	0.367
	Standard deviation	0.494	0.589	0.493	0.376

4.4.4. Cluster's characteristics

In this section, the cluster labels found in Section 4.4.3 are considered as the class labels. Using the features in Tables 3.1 and 3.2, the characteristics of the clusters are identified using CART with RPART package (Therneau et al. 2015). The decision tree for k=27 yields an accuracy of 0.738, and it is shown in Figure 4.7. Note that accuracy is calculated as the number of data points that are correctly classified divided by the total number of the data points (Han et al. 2012).

The decision tree indicates that clusters 1, 2, 5, 13, and 17 have 40, 10, 17, 1, and 4 members, respectively, and they can be characterized using the intermittency and long-term trend (moving average) features. These attributes can be interpreted as the sales pattern that is dominantly found in the corresponding clusters. Hence, the clusters are characterized by the features that show long-term behavior.

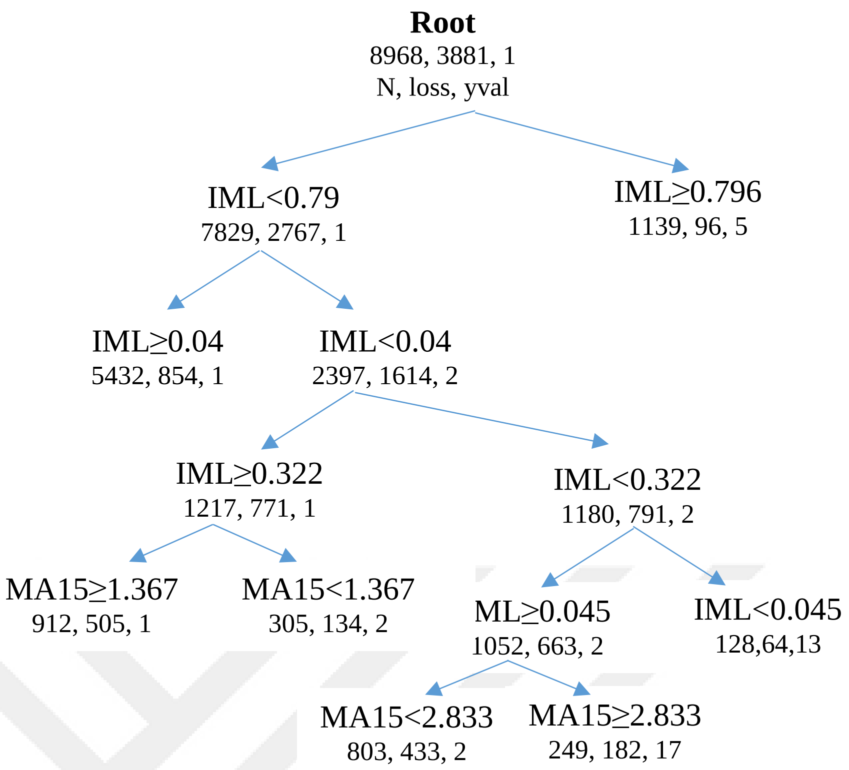


Figure 4.7. Decision tree for *k*=27 (N and loss denote the number of total points and the number of misclassified points, and yval denotes the cluster label.)

In Figure 4.7, the clusters can be characterized by multiple rules. For instance, cluster 1 can be identified using the long-term intermittency, i.e. IML<0.796, IML<0.407, IML≥0.323, and long-term trend, i.e. MA15≥1.367. Another way of characterizing cluster 1 is to use only long-term intermittency, i.e. IML<0.796 and IML≥0.407. In the figure, only cluster 5 includes data with high long-term intermittency (IML≥0.796) compared to the other clusters (IML<0.796). Meanwhile, cluster 13 is separated from the others with its low long-term intermittency (IML<0.045). Also, only clusters 2 and 17 are characterized by long-term trend, and both clusters have similar intermittency. The difference is that cluster 17 has a higher long-term trend (MA15≥2.833) than that of cluster 2 (MA15<2.833).

The decision trees for *k*=7 and *k*=16 are given in Appendix 6. Note that, for several clusters, no rules are generated because of the minimum number of observations in a node. Both decision trees have less number of leaf nodes compared to the one for *k*=27.

Also, all three decision trees are able to characterize five clusters using intermittency (IML) and trend (moving average) features. The difference is that the decision trees for k=7 and k=16 use mid-term trend (MA5) whereas the decision tree for k=27 uses long-term trend (MA15). Overall, the decision tree for k=27 yields more information about the characteristics of clusters.

4.4.5. Forecasting results

The performance of the proposed methodology is compared with other approaches. In this context, SVR, MARS+SVR, classification and regression tree (CART) and model tree (M5P) are used in the comparison. Note that CART and M5P are tree based approaches, and they perform feature selection implicity. The experiments are conducted using earth (Milborrow 2014), e1071 (Meyer et al. 2017), rpart (Therneau et al. 2015) and Rweka packages (Hornik et al. 2007).

In order to evaluate the impact of clustering on forecasting, the following model building alternatives are considered:

- Individual forecasting: 98 forecasting models are generated to predict the sales of each item. The forecasting models are built using SVR, SVR with MARS, CART and M5P, and these are abbreviated as S-SVR, S-MARS+SVR, S-CART, and S-M5P.

- Aggregate forecasting: For each month, the sales are summed over all products, and aggegate sales are generated. Then, a single SVR model is developed to predict the aggregate sales in the upcoming periods. Next, disaggregation is performed to determine the sales of each item. Two disaggregation methods are considered: i) for each product, the average market share over the planning horizon is used for disaggregation (abbreviated as A-SVR1), and ii) for each product, the total market share over the planning horizon is used for disaggregation (abbreviated as A-SVR2).

- Clustering-based forecasting: The methodology proposed in Section 3.2.6 is applied. That is, a forecasting model is built for each cluster prototype, so the number of forecasting models is equal to the number of clusters. The methodology is run with three different clustering results, i.e. k=7, k=16, and k=27, and two different cluster prototypes, i.e. medoid and DBA. SVR, SVR with MARS, CART and M5P are used in the development of the forecasting models, and these methods are denoted as C-SVR, C-MARS+SVR, C-CART, and C-M5P.

The performances of the approaches are compared in terms of forecasting error and complexity. MAPE, RMSE, MSE, and MAD are used for measuring the forecasting error. Complexity is considered as the number of features used in the forecasting models and the number of forecasting models.

For each setting, the average, maximum, minimum and standard deviation of the forecasting models are reported in Table 4.2. The results in Table 4.2 indicate that the best performances for average MAPE, RMSE, MSE and MAD are obtained for S-MARS+SVR in which all the items have a different forecasting model. Although its superiority in forecasting error, its complexity is high with a value of 1072, and 98 models need to be trained and tested.

Clustering improves the complexity significantly, whereas the forecasting errors worsen. In the clustering-based forecasting, the best average errors in terms of MAPE, RMSE, MSE and MAD are observed for C-MARS+SVR with k=27 and DBA. This model has an average MAPE of 31.66%, an average RMSE of 1.17, an average MSE of 3.24, an average MAD of 0.72, and a complexity value of 241. Thus, it results in 79.56% increase in MAPE, 31.64% increase in RMSE, 15.12% increase in MSE, 52.74% increase in MAD, and 77.52% decrease in complexity compared to S-MARS+SVR.

As the number of clusters decreases, the complexity decreases, e.g. for k=7, the complexity values of medoid and DBA methods are 80 and 69, respectively. However, the forecasting errors increase, e.g. for *k*=7 and medoid approach, MAPE, RMSE, MSE and MAD becomes 46.22%, 1.38, 3.91 and 0.92, respectively. Hence, it can be concluded that the sales patterns of the items are better identified for large number of clusters.

For *k*=27, DBA approach gives smaller errors compared to medoid approach, whereas, for *k*=7 and *k*=16, medoid yields smaller errors compared to DBA. So, DBA is a better approach for cluster prototype generation when the number of clusters is large.

Aggregate forecasting does not improve the error performance compared to the proposed approach. Also, CART and M5P result in higher error values both in individual and clustering-based forecasting.

Table 4.3 shows the best model for each model building alternative and forecasting error. Also, in Table 4.3, the relative performances of other models are compared with the best one. That is, the value of 0% shows the best model within the model building alternative and error performance. Positive values show the percentage deviation from the best model. The results indicate that MARS+SVR is the best model for almost all model building alternatives and forecasting errors. When *k*=7 and *k*=16, SVR with DBA is the best model in terms of MAPE. In all model building alternatives and forecasting errors, M5P and CART are significantly different from the best models. As these regression trees worsen the errors, they are not considered in the rest of the analysis.

Table 4.2. Comparison of the forecasting methods

Model building alternatives	Methods	Complexity	Average				Standard Deviation				Maximum				Minimum			
			MAPE (%)	RMSE	MSE	MAD	MAPE (%)	RMSE	MSE	MAD	MAPE (%)	RMSE	MSE	MAD	MAPE (%)	RMSE	MSE	MAD
individual (98-cluster)	S-SVR	2817	19.684	0.942	3.529	0.502	21.997	1.625	18.232	0.976	94.650	12.068	145.641	8.172	**0.000**	**0.000**	**0.000**	**0.000**
	S-MARS+SVR	1072	**17.630**	**0.885**	**2.811**	**0.474**	**16.435**	1.424	11.923	0.922	**60.887**	**8.791**	**77.286**	6.766	**0.000**	**0.000**	**0.000**	**0.000**
	S-M5P	205*	44.280	1.026	9.482	1.026	29.907	1.475	33.504	1.475	166.154	9.203	185.813	9.203	0.240	0.005	0.005	0.005
	S-CART	6*	44.697	1.092	9.572	1.092	24.753	1.681	39.691	1.681	138.846	11.656	247.719	11.656	0.721	0.010	0.010	0.010
aggregation (1-cluster)	A-SVR1	30	89.061	1.713	8.362	1.300	152.679	2.329	36.796	2.099	1119.123	16.887	285.173	15.462	9.398	0.485	0.236	0.236
	A-SVR2	30	83.094	1.749	4.881	1.322	8.625	**1.350**	**11.847**	**0.767**	101.002	9.610	92.361	**5.851**	64.540	0.933	0.870	0.841
7-cluster	C-SVR with k =7 and medoid	195	48.579	1.375	4.471	0.916	23.529	1.606	18.324	0.955	131.786	12.068	145.641	8.172	**0.000**	**0.000**	**0.000**	**0.000**
	C-SVR with k =7 and DBA	207	43.347	1.588	5.315	1.037	40.242	1.672	18.390	1.084	263.929	12.068	145.641	8.172	5.556	0.437	0.191	0.191
	C-MARS+SVR with k =7 and medoid	80	46.221	1.376	3.905	0.916	20.398	1.419	12.029	0.909	122.253	**8.791**	**77.286**	6.766	**0.000**	**0.000**	**0.000**	**0.000**
	C-MARS+SVR with k =7 and DBA	69	49.181	1.622	4.700	1.091	31.523	1.438	12.034	0.963	195.496	**8.791**	**77.286**	6.766	5.556	0.523	0.274	0.272
	C-M5P with k =7	92*	54.243	4.103	68.774	4.103	40.634	3.762	77.837	3.762	145.281	9.203	182.810	9.203	14.483	0.238	0.603	0.238
	C-CART with k =7	**5***	48.835	4.749	100.745	4.749	24.484	4.542	113.541	4.542	90.173	11.656	247.719	11.656	11.571	0.218	1.091	0.218
16-cluster	C-SVR with k =16 and medoid	464	44.505	1.262	4.088	0.820	21.880	1.580	18.303	0.922	100.000	12.068	145.641	8.172	**0.000**	**0.000**	**0.000**	**0.000**
	C-SVR with k =16 and DBA	477	32.783	1.354	4.355	0.816	19.517	1.588	18.273	0.936	100.000	12.068	145.641	8.172	**0.000**	**0.000**	**0.000**	**0.000**
	C-MARS+SVR with k =16 and medoid	143	42.026	1.239	3.440	0.809	18.289	1.380	11.985	0.867	100.000	**8.791**	**77.286**	6.766	**0.000**	**0.000**	**0.000**	**0.000**
	C-MARS+SVR with k =16 and DBA	132	38.213	1.390	3.869	0.885	32.382	1.392	11.933	0.900	195.496	**8.791**	**77.286**	6.766	**0.000**	**0.000**	**0.000**	**0.000**
	C-M5P with k =16	121*	66.906	2.893	35.175	2.893	40.535	2.789	59.829	2.789	166.154	9.203	182.810	9.203	14.483	0.238	0.603	0.238
	C-CART with k =16	**5***	63.810	3.166	48.368	3.166	30.826	3.345	88.251	3.345	138.846	11.656	247.719	11.656	11.571	0.218	1.091	0.218
27-cluster	C-SVR with k =27 and medoid	781	39.671	1.248	3.991	0.791	23.753	1.560	18.255	0.915	100.000	12.068	145.641	8.172	**0.000**	**0.000**	**0.000**	**0.000**
	C-SVR with k =27 and DBA	781	32.778	1.220	3.957	0.746	21.844	1.571	18.260	0.930	94.650	12.068	145.641	8.172	**0.000**	**0.000**	**0.000**	**0.000**
	C-MARS+SVR with k =27 and medoid	254	37.321	1.185	3.248	0.754	18.907	1.358	11.903	0.868	100.000	**8.791**	**77.286**	6.766	**0.000**	**0.000**	**0.000**	**0.000**
	C-MARS+SVR with k =27 and DBA	241	31.656	1.165	3.236	0.724	18.640	1.370	11.907	0.881	104.968	**8.791**	**77.286**	6.766	**0.000**	**0.000**	**0.000**	**0.000**
	C-M5P with k =27	31*	67.182	2.235	29.165	2.235	38.009	2.332	57.420	2.332	166.154	9.203	185.813	9.203	8.975	0.149	0.236	0.149
	C-CART with k =27	7*	60.797	2.355	31.151	2.355	28.648	2.783	71.158	2.783	138.846	11.656	247.719	11.656	10.278	0.172	0.321	0.172

*both M5P and CART report the number of rules

Table 4.3. Percentage of error increase compared to the best method

Model building alternatives	Methods	% error increase compared to the best method			
		MAPE	RMSE	MSE	MAD
individual (98-cluster)	S-SVR	11.65	6.38	25.53	5.90
	S-MARS+SVR	**0.00**	**0.00**	**0.00**	**0.00**
	S-M5P	151.16	15.94	237.31	116.53
	S-CART	153.52	23.39	240.52	130.44
7-cluster	C-SVR with *k*=7 and medoid	12.07	**0.00**	14.50	**-0.10**
	C-SVR with *k*=7 and DBA	**0.00**	15.44	36.12	13.20
	C-MARS+SVR with *k*=7 and medoid	6.63	**0.02**	**0.00**	**0.00**
	C-MARS+SVR with *k*=7 and DBA	13.46	17.95	20.37	19.00
	C-M5P with *k*=7	25.14	198.32	1661.34	347.72
	C-CART with *k*=7	12.66	245.32	2480.13	418.26
16-cluster	C-SVR with *k*=16 and medoid	35.76	1.80	18.86	1.37
	C-SVR with *k*=16 and DBA	**0.00**	9.23	26.60	0.82
	C-MARS+SVR with *k*=16 and medoid	28.20	**0.00**	**0.00**	**0.00**
	C-MARS+SVR with *k*=16 and DBA	16.56	12.13	12.49	9.29
	C-M5P with *k*=16	104.09	133.43	922.62	257.51
	C-CART with *k*=16	94.64	155.43	1306.20	291.20
27-cluster	C-SVR with *k*=27 and medoid	25.32	7.08	23.32	9.14
	C-SVR with *k*=27 and DBA	3.54	4.73	22.28	2.93
	C-MARS+SVR with *k*=27 and medoid	17.90	1.69	0.35	4.10
	C-MARS+SVR with *k*=27 and DBA	**0.00**	**0.00**	**0.00**	**0.00**
	C-M5P with *k*=27	112.22	91.74	801.23	208.48
	C-CART with *k*=27	92.05	102.05	862.60	225.06

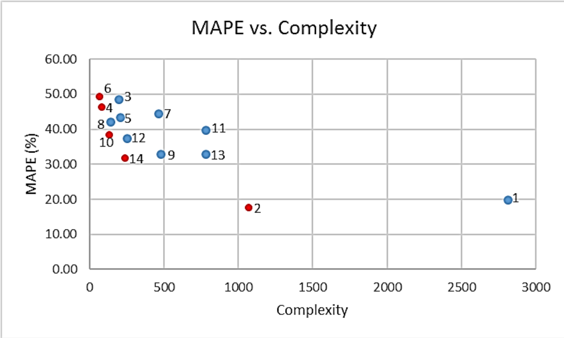
The aim is to obtain an accurate forecasting model with low complexity. Hence, two criteria are used to evaluate the forecasting models. The first criterion is the forecasting error, i.e. MAPE, RMSE, MSE, or MAD. The second criterion is the complexity, i.e. the number of forecasting models and the number of features used in the model. Hence, the best forecasting model is selected within the multiple criteria (attribute) decision making framework. In this context, a solution is *dominated* if there are other solutions that are better than it in at least one criterion and as good as it in other criteria (Yoon and Hwang 1995). If a solution is not dominated by other solutions, it is called *non-dominated solution* (Yoon and Hwang 1995).

Figure 4.8 presents the non-dominated solutions with respect to the complexity and error performances. From the graphs, S-MARS+SVR, C-MARS+SVR with k=27 and DBA, C-MARS+SVR with k=7 and medoid, and C-MARS+SVR with k=7 and DBA are the non-dominated solutions based on all error and complexity pairs. Additionally, C-MARS+SVR with k=16 and DBA is a non-dominated solution for all criteria pairs, except RMSE and complexity criteria. Also, C-MARS+SVR with k=16 and medoid is a non-dominated solution for all criteria pairs, except MAPE and complexity criteria.

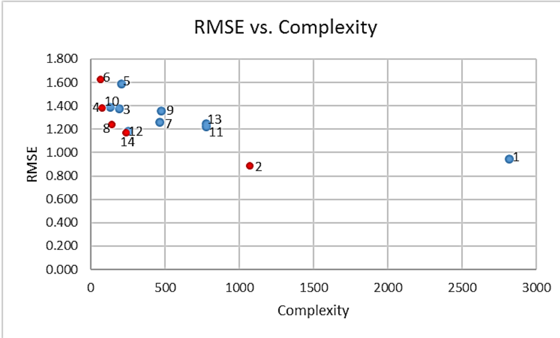
The difference between the non-dominated solutions are tested using the non-parametric tests, namely sign test and Wilcoxon signed rank test. Appendix 7 shows the Minitab outputs of Wilcoxon signed rank test. The results indicate that the non-dominated solutions are significantly different from each other (p-value < 0.05).

To sum up, Table 4.4 summarizes the relative performance of the six non-dominated solutions with respect to the base case, S-MARS+SVR. Note that base case is selected as it has the minimum forecasting error. The percentages of increase in the forecasting errors and decrease in the complexity with respect to the base case are provided in Table 4.4. Among the non-dominated solutions, C-MARS+SVR with k=27 and DBA provides a reduction of 77.52% in the complexity, whereas the forecasting errors increase with percentages of 79.55%, 31.65%, 15.12%, 52.82%. For the other non-dominated solutions, the marginal improvement in the complexity is less than the marginal increase in the forecasting errors.

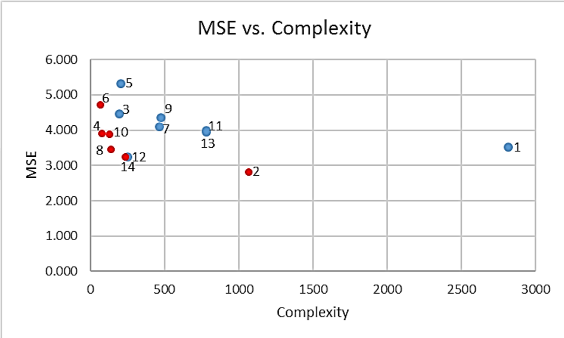
An interesting observation is that the highest error increase is for MAPE. This is due to the limitation of MAPE in the intermittent data. That is, if the actual sales of one unit is predicted as zero in a highly intermittent item, MAPE becomes 100%. Hence, it is more fair to use RMSE, MSE, and MAD to measure forecasting errors in intermittent data.



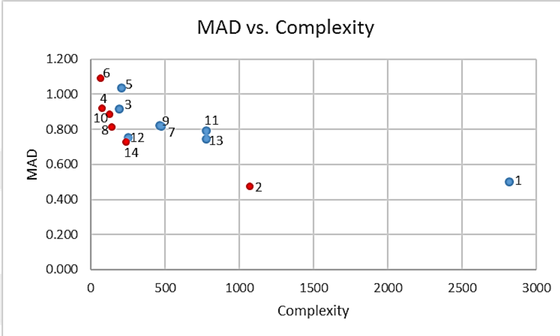
(a)



(b)



(c)



(d)

Legend:

1. Pure SVR for individual forecasting
2. SVR+MARS for individual forecasting
3. SVR+Medoid for 7-clusters
4. SVR+MARS-Medoid for 7-clusters
5. SVR+DBA for 7-clusters
6. SVR+MARS-DBA for 7-clusters
7. SVR+Medoid for 16-clusters
8. SVR+MARS-Medoid for 16-clusters
9. SVR+DBA for 16 clusters
10. SVR+MARS-DBA for 16 clusters
11. SVR+Medoid for 27 clusters
12. SVR+MARS-Medoid for 27 clusters
13. SVR+DBA for 27 clusters
14. SVR+MARS-DBA for 27 clusters

Figure 4.8. Non-dominated solutions with respect to the forecasting error and complexity

Table 4.4. Relative comparison of the nondominated solutions

Methods	% change				
	complexity	**MAPE**	**RMSE**	**MSE**	**MAD**
S-MARS+SVR	0.00	0.00	0.00	0.00	0.00
C-MARS+SVR with k=27 and DBA	-77.52	79.55	31.65	15.12	52.82
C-MARS+SVR with k=16 and DBA	-87.69	116.75	100.00	37.64	86.60
C-MARS+SVR with k=16 and medoid	-86.66	100.00	40.02	22.36	70.73
C-MARS+SVR with k=7 and medoid	-92.54	162.17	55.40	38.90	93.33
C-MARS+SVR with k=7 and DBA	-93.56	178.96	83.25	67.20	130.07

4.4.6. Results of inventory performance

In this section, the inventory performance of C-MARS+SVR with k=27 and DBA is evaluated as explained in Section 3.2.7.

Two items are selected as a sample. One item (item 26) has a low intermittency level (0.29), whereas the other (item 11) has a high intermittency level (0.77), and both items have time series with lengths of 228. For each item, 80% of the sales data are used for training and testing the forecasting model. The remaining 20% is used for the evaluation of the inventory performance.

In the inventory evaluation procedure, alternative scenarios are examined, and the best scenario that couples with proposed forecasting methodology is determined. In this context, the following twelve alternative scenarios are studied:

Scenario 1: Zero safety stock and LFL strategy
Scenario 2: Zero safety stock and fixed 3-period strategy
Scenario 3: 1 unit of safety stock and LFL strategy
Scenario 4: 1 unit of safety stock and fixed 3-period strategy
Scenario 5: 2 units of safety stock and LFL strategy
Scenario 6: 2 units of safety stock and fixed 3-period strategy
Scenario 7: 3 units of safety stock and LFL strategy
Scenario 8: 3 units of safety stock and fixed 3-period strategy
Scenario 9: 4 units of safety stock and LFL strategy
Scenario 10: 4 units of safety stock and fixed 3-period strategy
Scenario 11: 5 units of safety stock and LFL strategy
Scenario 12: 5 units of safety stock and fixed 3-period strategy

Each alternative scenario is evaluated in terms of total cost, average inventory level, backorder level, and IT. In the forklift distributor company, lead time is three months. The inventory holding cost factor is 11% per year/unit. The unit purchase price is 10,840$/unit. The fixed order cost is 13,500$/order. The unit backordering cost is 434$

per month/unit. Initial inventory levels are varied between zero and three in order to analyze the impact of starting conditions.

Table 4.5 summarizes the inventory performances of the two items. The details of the inventory performance evaluation with scenario 1 is given in Appendix 8. Note that, in the table, the “Run-Out Time” column shows the waiting time of inventory in the warehouse. The aim is to minimize this value.

Table 4.5. Evaluation of the inventory performance

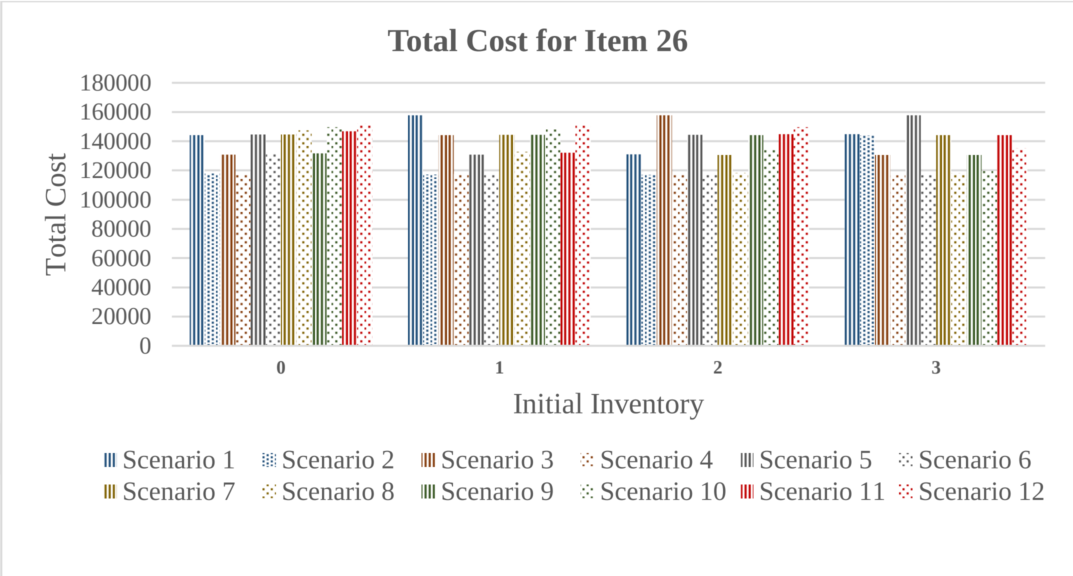
No	Item	Intermittency Level	Scenario	Initial inventory	Total Cost	Total Backorder	Average Inventory	Annual Sales	IT	Run-Out Time
1	26	0.29	1	2	131117	7	2.58	27.79	20.31	0.43
2	26	0.29	2	2	117054	6	3.89	27.79	10.46	0.83
3	26	0.29	3	3	130706	5	3.05	27.79	15.09	0.57
4	26	0.29	4	3	116841	3	4.47	27.79	8.59	1.01
5	26	0.29	5	2	144328	6	3.58	27.79	11.73	0.74
6	26	0.29	6	3	116925	2	5.21	27.79	7.04	1.23
7	26	0.29	7	3	144115	4	4.16	27.79	9.43	0.92
8	26	0.29	8	3	118075	2	6.05	27.79	5.83	1.49
9	26	0.29	9	3	130600	4	4.84	27.79	7.65	1.13
10	26	0.29	10	3	119621	1	7.11	27.79	4.78	1.81
11	26	0.29	11	3	144283	2	5.63	27.79	6.32	1.37
12	26	0.29	12	3	135299	1	8.26	27.79	3.98	2.17
13	11	0.77	1	1	44041	2	0.73	0.92	1.31	6.62
14	11	0.77	2	1	44041	2	0.73	0.92	1.31	6.62
15	11	0.77	3	1	44201	1	0.88	0.92	1.09	7.94
16	11	0.77	4	1	44201	1	0.88	0.92	1.09	7.94
17	11	0.77	5	2	47430	0	1.83	0.92	0.51	16.85
18	11	0.77	6	2	47430	0	1.83	0.92	0.51	16.85
19	11	0.77	7	0	44226	0	4.50	0.92	0.21	41.53
20	11	0.77	8	0	44226	0	4.50	0.92	0.21	41.53
21	11	0.77	9	1	48087	0	5.50	0.92	0.17	50.92
22	11	0.77	10	1	48087	0	5.50	0.92	0.17	50.92
23	11	0.77	11	2	51948	0	6.50	0.92	0.14	60.31
24	11	0.77	12	2	51948	0	6.50	0.92	0.14	60.31

The results indicate that the item with the low intermittency level (item 26) has the minimum total cost in scenario 4 and the maximum IT in scenario 1. Scenario 1 indicates that no safety stock together with LFL achieves the maximum IT. Hence, the company’s orders should be equal to the forecasted demand for one period. However, the total number of backorders is the highest with a value of seven in this strategy, and the total cost is high. On the other hand, scenario 4 suggests keeping one unit safety stock and using fixed 3-period inventory policy. This scenario provides the minimum total cost, and reduces the number of backorders to three. As the unit inventory holding cost is smaller than the unit backordering cost, the company favors holding inventory compared to backordering. Compared to scenario 4, scenario 1 provides an improvement of 57.72% in the IT and it causes an increase of 10.88% in the total cost. The company’s main goal is to maximize the IT value. That is, the company prefers selling the item as soon as possible without keeping it in the warehouse for a long time. The backordering case is not important due to the loyalty of the customers. For these reasons, scenario 1 is more appropriate to the company for the item with the low intermittency level.

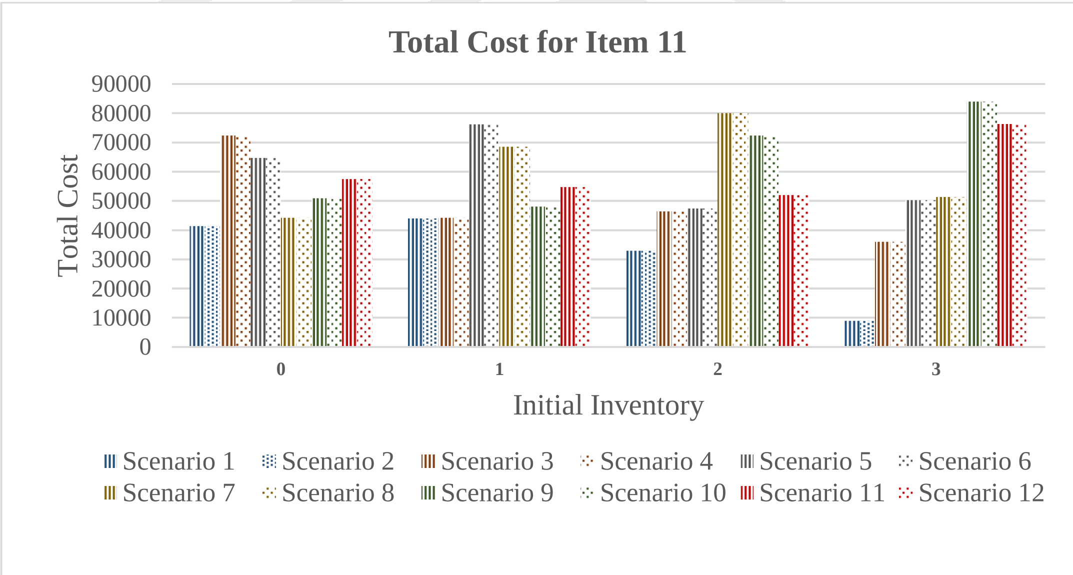
For the item with high intermittency level (item 11), scenarios 1 and 2 yield the best performance results in terms of total cost and IT. That is, either LFL or fixed 3-period strategy can be coupled with zero safety stock level. The zero safety stock level makes sense as the high intermittency results in many zero values of demand.

The robustness of the scenarios to initial inventory level is also evaluated. Varying the initial inventory levels, the total cost and IT values of different scenarios are presented in Figure 4.9 and Figure 4.10, respectively. Figure 4.9 indicates that, in terms of total cost, the best scenarios for the items having low and high intermittency levels (items 26 and 11) are scenarios 4 and 1, respectively. Hence, the initial inventory conditions do not change the safety stock level and lot sizing decisions for both items. In Figure 4.10, in terms of IT, the best scenario is scenario 1 for both items. Thus, scenario 1 which includes LFL strategy is more robust to the initial inventory level compared to scenario 2 which includes fixed 3-period strategy.

As a summary, the proposed forecasting methodology should couple with the zero safety stock and LFL strategy for items having either low or high intermittency, as the company is interested in maximizing the IT. For the item with high intermittency level, this strategy results in the minimum total cost as well. For the item with the low intermittency level, this strategy causes an increase of 10.88% in the total cost. Another advantage of the proposed approach is its robustness to initial inventory levels. Also, the company holds safety stocks to satisfy the customer demand currently, whereas the proposed forecasting methodology helps the company eliminate the safety stocks.

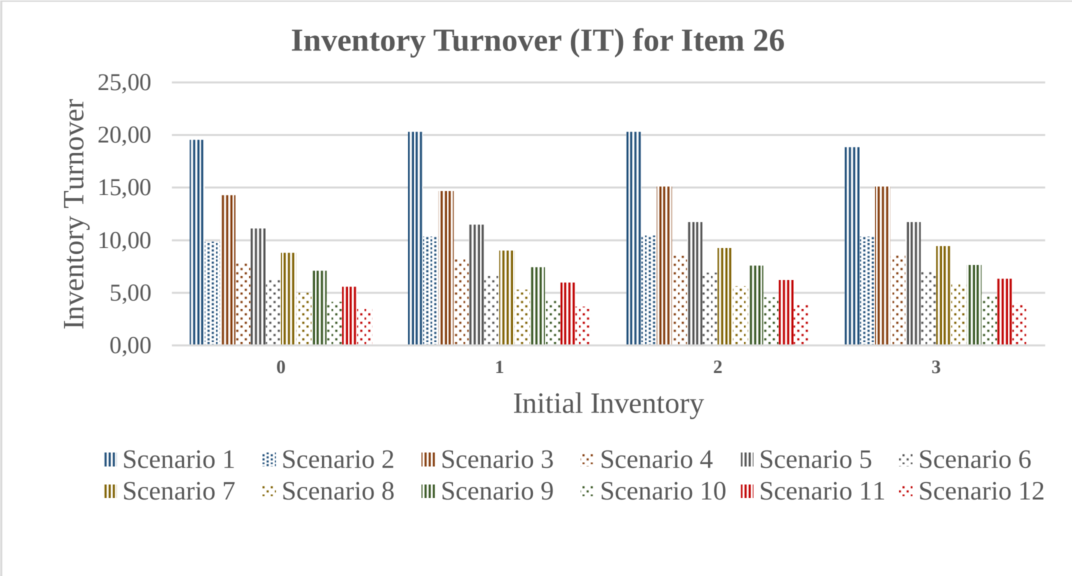


(a)

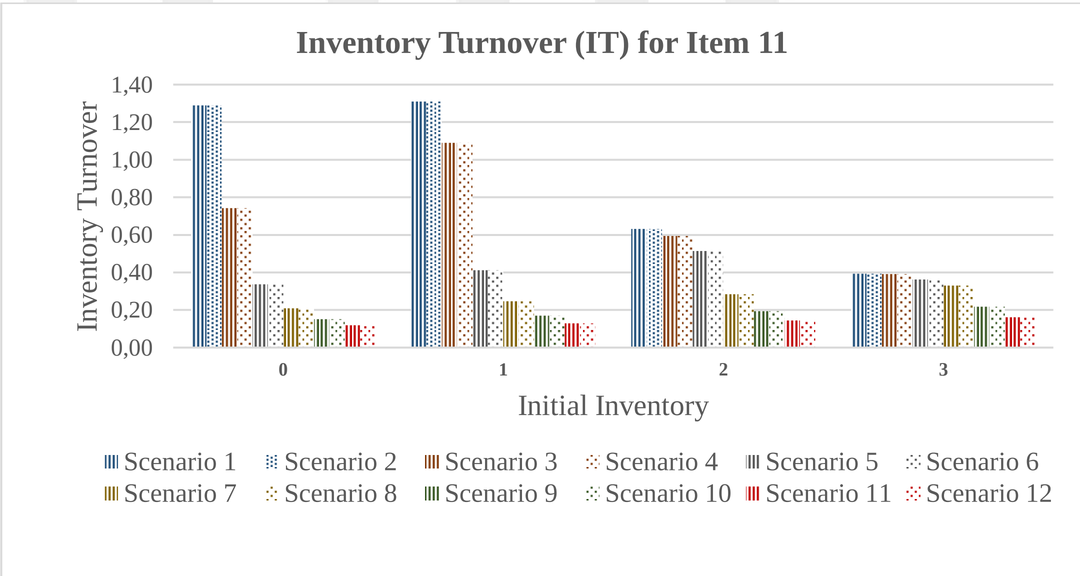


(b)

Figure 4.9. Comparison of the scenarios in terms of total cost with different initial inventory levels for (a) item 26 and (b) item 11



(a)



(b)

Figure 4.10. Comparison of the scenarios in terms of IT with different initial inventory levels for (a) item 26 and (b) item 11

5. DISCUSSIONS AND CONCLUSION

5.1. Discussion

The proposed forecasting methodology is compared with the current forecasting method used in the company. Currently, the company applies the aggregate approach to minimize the efforts in forecasting. As shown in Section 4.4.5, the aggregate forecasting has a single forecasting model with 30 features, so its complexity is low. However, its error values are quite high, i.e. MAPE, MSE, RMSE and MAD values are 89.1%, 1.713, 8.362, and 1.3, respectively.

The use of the proposed approach ensures flexibility to the decision maker (sales manager) in balancing accuracy and complexity. That is, it yields a number of non-dominated solutions, and the sales manager can select one of them. For example, one of the non-dominated solutions improves the forecasting errors such that i.e. MAPE, MSE, RMSE and MAD values are 31.7%, 1.165, 3.236, and 0.724, respectively. In this case, the complexity becomes 241.

The proposed methodology facilitates the implementation and interpretation of forecasting processes in the company. For example, the decision tree proposed in Section 3.2.4 shows that the clusters differ in terms of trend and intermittency. Another advantage of the proposed approach is that no safety stock policy is favored for the items with low and high intermittency levels.

As a summary, the proposed approach is in line with the company's aim, it ensures reasonable accuracy with less complexity. It provides insights about the sales patterns of the different items, and improves the inventory management system.

5.2. Conclusion

Forecasting methods that ensure high accuracy with less complexity are valuable for the companies with high product variety, as effective forecasting mechanisms reduce the inventory costs and increases the customer satisfaction. Motivated by this, a data mining-based forecasting methodology, that combines clustering, feature extraction, feature selection, and classification, is proposed.

The proposed methodology addresses the time series sequences with unequal lengths and intermittency. In order to reduce the number of forecasting models, items with similar sales patterns are clustered using DTW distance. Then, each cluster is represented by the cluster prototype, and these cluster prototypes are used for forecasting. Thus, instead of using all products, a simplication is achieved. Also, cluster characteristics are extracted using classification methods.

Next, novel features are introduced to handle intermittent data. Also, several features that characterize trend, volatility, and growth, are adopted from the literature. MARS is used to select the useful features for forecasting the sales, and a forecasting model is built for each cluster using SVR. The results indicate the superiority of combining MARS and SVR compared to the other approaches. The proposed approach gives reasonable accuracy with low complexity. Also, the inventory performance evaluation of the proposed forecasting approach indicates that it helps the company reduce the safety stock requirements and total inventory cost.

The proposed approach can be applicable to intermittent data, and it provides flexibility to achieve low complexity and high accuracy. The decision maker can select an appropriate forecasting model according to his/her aims.

The proposed methodology has a broad impact on business. That is, it can be used to forecast the sales of the new products. Moreover, a wide variety of companies such as retailers of fast fashion can benefit from the proposed methodology.

Sales data may include noise and outliers, so future studies can focus on the integration of mechanisms to handle outlier and noise. Also, future studies may include the development of error measures for intermittent data.

REFERENCES

Agrawal, R., Faloutsos, C., and Swami, A. 1993. Efficient similarity search in sequence databases. The Fourth International Conference Foundations of Data Organization and Algorithms, 69-84.

Altay, N., Rudisill, F., and Litteral, L. A. 2008. Adapting Wright's modification of Holt's method to forecasting intermittent demand. *International Journal of Production Economics*, 111(2): 389–408.

Anh, D. T., and Thanh, L. H. 2015. An efficient implementation of k-means clustering for time series data with DTW distance. *International Journal of Business Intelligence and Data Mining*, 10(3): 213–232.

Austin, P. C., Tu, J. V., Ho, J. E., Levy, D., and Lee, D. S. 2013. Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes. *Journal of clinical epidemiology*, 66(4): 398-407.

Bala, P. K. 2010. Purchase-driven classification for improved forecasting in spare parts inventory replenishment. *International Journal of Computer Applications*, 10(9): 40–45.

Bala, P.K. 2012. Improving inventory performance with clustering based demand forecasts. *Journal of Modelling in Management*, 7(1): 23-37.

Bao, Y., Wang, W., and Zhang, J. 2004. Forecasting intermittent demand by SVMs regression. IEEE International Conference on Systems, Man and Cybernetics, 1: 461–466.

Bao, Y., Wang, W., and Zou, H. 2005. SVR-based method forecasting intermittent demand for service parts inventories. Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, 604–613.

Barrack, D., Goulding, J., Hopcraft, K., Preston, S., and Smith, G. 2015. AMP: a new time-frequency feature extraction method for intermittent time-series data. Retrieved from http://arxiv.org/abs/1507.05455

Bermúdez, J. D., Segura, J. V., and Vercher, E. 2006. Improving demand forecasting accuracy using nonlinear programming software. *Journal of the Operational Research Society*, 57(1): 94–100.

Berndt, D., and Clifford, J. 1994. Using dynamic time warping to find patterns in time series. Workshop on Knowledge Knowledge Discovery in Databases, 398: 359–370.

Biscarri, F., Monedero, I., García, A., Guerrero, J. I., and León, C. 2017. Electricity clustering framework for automatic classification of customer loads. *Expert Systems with Applications*, 86: 54–63.

Box, G.E., Jenkins, G.M. 1976. Time series analysis: forecasting and control. Holden-Day.

Boylan, J. E., and Syntetos, A. A. 2010. Spare parts management: a review of forecasting research and extensions. *IMA Journal of Management Mathematics*, 21(3): 227–237.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. J. 1984. Classification and regression tree. Wadsworth International Group, Belmont, CA.

Brockwell, P., and Davis, R. 2002. Introduction to time series and forecasting. Technometrics.

Brown, R. G. 1959. Statistical forecasting for inventory control. McGraw/Hill.

Chakrabarti, K., Keogh, E., Mehrotra, S., and Pazzani, M. 2002. Locally adaptive dimensionality reduction for indexing large time series databases. *ACM Transactions on Database Systems*, 27(2): 188–228.
Chen, I. F., and Lu, C. J. 2017. Sales forecasting by combining clustering and machine-learning techniques for computer retailing. *Neural Computing and Applications*, 28(9): 2633–2647.
Chen, Q., Hu, G., Gu, F., and Xiang, P. 2012. Learning optimal warping window size of DTW for time series classification. 11th International Conference on Information Science, Signal Processing and Their Applications, 1272–1277.
Choi, T. M., Hui, C. L., Liu, N., Ng, S. F., and Yu, Y. 2014. Fast fashion sales forecasting with limited data and time. *Decision Support Systems*, 59(1): 84–92.
Coyle, J. J., Bardi, E. J., and Langley, C. J. 2003. The Management of business logistics: a supply chain perspective. South-Western Thomson Learning, Ohio.
Columbus, L. 2015. Roundup of analytics, big data and business intelligence and market estimates.https://www.forbes.com/sites/louiscolumbus/2015/05/25/roundup-of-analytics-big-data-business-intelligence-forecasts-and-market-estimates-2015/#420fc96444f3-(Downloaded: 05/05/2017).
Croston, J. D. 1972. Forecasting and stock control for intermittent demands. Operational Research Quarterly, 289-303.
Dai, W., Chuang, Y.-Y., and Lu, C.-J. 2015. A Clustering-based sales forecasting scheme using support vector regression for computer server. *Procedia Manufacturing*, 2: 82–86.
Das, S., and Padhy, S. 2012. Support vector machines for prediction of futures prices in Indian stock market. *International Journal of Computer Applications*, 41(3): 22–26.
Duch, Wł., Setiono, R., and Zurada, J. M. 2004. Computational intelligence methods for rule-based data understanding. Proceedings of the IEEE, 92(5): 771–805.
Friedman, J. H. 1991. Multivariate adaptive regression splines. *The Annals of Statistics*, 19(1): 1–67.
Fu, A. W. C., Keogh, E., Lau, L. Y. H., Ratanamahatana, C. A., and Wong, R. C. W. 2008. Scaling and time warping in time series querying. *VLDB Journal*, 17(4): 899–921.
Gilan, S., Bahrami, J. H., and Ramezanianpour, A. A. 2012. hybrid support vector regression - particle swarm optimization for prediction of compressive strength and rcpt of concretes containing metakaolin. *Construction and Building Materials*, 34: 321–329.
Giorgino, T. 2009. Computing and visualizing dynamic time warping alignments in R : The dtw Package. *Journal of Statistical Software*, 31(7): 1–24.
Górecki, T. 2014. Using derivatives in a longest common subsequence dissimilarity measure for time series classification. *Pattern Recognition Letters*, 45(1): 99–105.
Gorunescu, F. 2011. Data mining: concepts, models, and techniques. Springer.
Greche, L., Jazouli, M., Es-Sbai, N., Majda, A., and Zarghili, A. 2017. Comparison between Euclidean and Manhattan distance measure for facial expressions classification. 2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems, 2–5.
Han, J., Kamber, M., and Pei, J. 2012. Data mining: concepts and techniques. Morgan Kaufmann, San Francisco, CA.
Hautamaki, V., Nykanen, P., and Franti, P. 2008. Time-series clustering by approximate prototypes. 19th International Conference on Pattern Recognition, 1–4.

He, W., Zha, S., and Li, L. 2013. Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, 33(3): 464-472.
Hornik, K., Karatzoglou, D. M., Zeileis, A., and Hornik, M. K. 2007. The rweka package.
Hua, Z., and Zhang, B. 2006. A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts. *Applied Mathematics and Computation*, 181(2): 1035–1048.
Huber, J., Gossmann, A., and Stuckenschmidt, H. 2017. Cluster-based hierarchical demand forecasting for perishable goods. *Expert Systems with Applications*, 76: 140–151.
Hyndman, R. J., Lee, A. J., and Wang, E. 2014. Fast computation of reconciled forecasts for hierarchical and grouped time series. *Computational Statistics & Data Analysis*, *97*: 16-32.
IBM Research. 2011. The role of data mining in business optimization. IBM Corporation.
Itakura, F. 1975. Minimum prediction residual principle applied to speech recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 23(1): 67–72.
Jain, A. K. 2010. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8): 651–666.
Jeong, Y. S., Jeong, M. K., and Omitaomu, O. A. 2011. Weighted dynamic time warping for time series classification. *Pattern Recognition*, 44(9): 2231–2240.
Johnston, F. R., and Boylan, J. E. 1996. Forecasting for items with intermittent demand. *Journal of the Operational Research Society*, 47(1): 113–121.
Keogh, E. 1997. A fast and robust method for pattern matching in time series databases. Proceedings of WUSS, 97(1): 99.
Keogh, E. J., and Pazzani, M. J. 2000. Scaling up dynamic time warping for datamining applications. Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 285–289.
KernelSVM. 2018. Support vector machine regression. http://kernelsvm.tripod.com- (Downloaded: 12/02/2018).
Kirshners, A. K., Parshutin, S. V., and Borisov, A. N. 2010. Combining clustering and a decision tree classifier in a forecasting task. *Automatic Control and Computer Sciences*, 44(3): 124–132.
Kumar, V., and Rathi, N. 2011. Knowledge discovery from database Using an integration of clustering and classification. *International Journal of Advanced Computer Science and Applications*, 2(3): 29–33.
Kuo, R. J., and Li, P. S. 2016. Taiwanese export trade forecasting using firefly algorithm based K-means algorithm and SVR with wavelet transform. *Computers and Industrial Engineering*, 99: 153–161.
Lai, R. K., Fan, C.-Y., Huang, W.-H., and Chang, P.-C. 2009. Evolving and clustering fuzzy decision tree for financial time series data forecasting. *Expert Systems with Applications*, 36(2): 3761–3773.
Lei, H., and Govindaraju, V. 2004. Regression time warping for similarity measure of sequence. The Fourth International Conference of IEEE, 826-830.
Levis, A. A., and Papageorgiou, L. G. 2005. Customer demand forecasting via support vector regression analysis. *Chemical Engineering Research and Design*, 83(8 A): 1009–1018.

Liao, W. T. 2005. Clustering of time series data - a survey. *Pattern Recognition*, 38(11): 1857–1874.

Lines, J., and Bagnall, A. 2015. Time series classification with ensembles of elastic distance measures. *Data Mining and Knowledge Discovery*, 29(3): 565–592.

Loureiro, A., Torgo, L., and Soares, C. 2004. Outlier detection using clustering methods: a data cleaning application. Proceedings of KDNet Symposium on Knowledge-based Systems for the Public Sector, Bonn, Germany.

Lu, C.-J. 2014. Sales forecasting of computer products based on variable selection scheme and support vector regression. *Neurocomputing*, 128: 491–499.

Lu, C. J., and Kao, L. J. 2016. A clustering-based sales forecasting scheme by using extreme learning machine and ensembling linkage methods with applications to computer server. *Engineering Applications of Artificial Intelligence*, 55: 231–238.

Lu, C. J., Lee, T. S., and Chiu, C. C. 2009. Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*, 47(2): 115–125.

Lu, C. J., Lee, T. S., and Lian, C. M. 2012. Sales forecasting for computer wholesalers: A comparison of multivariate adaptive regression splines and artificial neural networks. *Decision Support Systems*, 54(1): 584–596.

Maimon, O., and Rokach, L. 2005. The data mining and knowledge discovery handbook. Springer, New York.

Mentzer, J. T., and Moon, M. A. 2004. Sales forecasting management: a demand management approach. Sage.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C. C., Lin, C. C., and Meyer, M. D. 2017. Package 'e1071'.

Milborrow, S. 2014. Notes on the earth package.

Moon, M. A., Mentzer, J. T., and Smith, C. D. 2003. Conducting a sales forecasting audit. *International Journal of Forecasting*, 19(1): 5–25.

Murray, P. W., Agard, B., and Barajas, M. A. 2015. Forecasting supply chain demand by clustering customers. *IFAC-PapersOnLine*, 28(3): 1834–1839.

Murray, P. W., Agard, B., and Barajas, M. A. 2017. Market segmentation through data mining: A method to extract behaviors from a noisy data set. *Computers & Industrial Engineering*, 109: 233–252.

Murugavel, P, and Punithavalli, M. 2011. Improved hybrid clustering and distance-based technique for outlier removal. *International Journal on Computer Science and Engineering*, 3(1): 333–339.

Nalbantov, G., Groenen, P. J., and Bioch, J. C. 2005. Support vector regression basics. *Medium Econometrische Toepassingen*, 13(1): 16-19.

O'Marah, K., John, G., Blake, B., and Manenti, P. 2014. The chief supply chain officer report 2014. SCM World, 34.

Ou, P., and Wang, H. 2009. Prediction of stock market index movement by ten data mining techniques. *Modern Applied Science*, 3(12): 28–42.

Petitjean, F., Ketterlin, A., and Gançarski, P. 2011. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition*, 44(3): 678–693.

Quinlan, J. R. 1986. Induction of decision trees. *Machine Learning*, 1(1): 81–106.

Quinlan, J. R. 1993. C4. 5: programs for machine learning. Morgan Kaufmann, San Mateo, CA.

R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Radhakrishnan, B., Shineraj, G., and Anver M. K. M. 2013. Application of data mining in marketing. *International Journal of Computer Science and Network*, 2(5): 2277-5420.

Reid, R. D., and Sanders, N. R. 2005. Operations management: an integrated approach. John Wiley, Hoboken, NJ.

Sayad, S. 2018. Hierarchical clustering.
http://www.saedsayad.com/clustering_hierarchical.htm-(Downloaded: 12/02/2018).

Sakoe, H., and Chiba, S. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1): 43–49.

Sharma, M. 2014. Data Mining : a literature survey. *International Journal of Emerging Research in Management &Technology*, 9359(2): 1–4.

Snyder, R. 2002. Forecasting sales of slow and fast moving inventories. *European Journal of Operational Research*, 140(3): 684–699.

So, K. C., and Zheng, X. 2003. Impact of supplier's lead time and forecast demand updating on retailer's order quantity variability in a two-level supply chain. *International Journal of Production Economics*, 86(2): 169–179.

Soheily-Khah, S., Douzal-Chouakria, A., and Gaussier, E. 2015. Progressive and iterative approaches for time series averaging. CEUR Workshop Proceedings, 1425: 123–129.

Syntetos, A. A., and Boylan, J. E. 2001. On the bias of intermittent demand estimates. *International Journal of Production Economics*, 71(1–3): 457–466.

Tan, P.-N., Steinbach, M., and Kumar, V. 2006. Classification: basic concepts, decision trees, and model evaluation. *Introduction to Data Mining*, 67(17): 145–205.

Taylan, P., Weber, G. W., and Özkurt, F. Y. 2010. A new approach to multivariate adaptive regression splines by using Tikhonov regularization and continuous optimization. *Top*, 18(2): 377-395.

Therneau, T., Atkinson, B., Ripley, B., and Ripley, M. B. 2015. Package 'rpart'.

Thissen, U., Van Brakel, R., De Weijer, A. P., Melssen, W. J., and Buydens, L. M. C. 2003. Using support vector machines for time series prediction. *Chemometrics and Intelligent Laboratory Systems*, 69(1–2): 35–49.

Thomassey, S. 2010. Sales forecasts in clothing industry: the key success factor of the supply chain management. *International Journal of Production Economics*, 128(2): 470–483.

Thomassey, S., and Fiordaliso, A. 2006. A hybrid sales forecasting system based on clustering and decision trees. *Decision Support Systems*, 42(1): 408–421.

Thomassey, S., and Happiette, M. 2007. A neural clustering and classification system for sales forecasting of new apparel items. *Applied Soft Computing Journal*, 7(4): 1177–1187.

Tormene, P., Giorgino, T., Quaglini, S., and Stefanelli, M. 2009. Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. *Artificial Intelligence in Medicine*, 45(1): 11–34.

Tsai, C. Y., Tsai, C. Y., and Huang, P. W. 2009. An association clustering algorithm for can-order policies in the joint replenishment problem. *International Journal of Production Economics*, 117(1): 30–41.

Tsai, M. F., and Yu, S. S. 2016. Distance metric based oversampling method for bioinformatics and performance evaluation. *Journal of Medical Systems*, 40(7).

Vapnik, V. N. 1998. Statistical learning theory. *Adaptive and Learning Systems for Signal Processing, Communications and Control*, 2: 1–740.

Wang, J., Li, L., Niu, D., and Tan, Z. 2012. An annual load forecasting model based on support vector regression with differential evolution algorithm. *Applied Energy*, 94: 65–70.

Wei, J. T., Lee, M. C., Chen, H. K., and 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.

Winters, P. R. 1960. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3): 324–342.

Wisner, J. D., Tan, K. C., and Leong, G. K. 2014. Principles of supply chain management: A balanced approach. Cengage Learning.

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. 2016. Data mining: practical machine learning tools and techniques. Morgan Kaufmann.

Wu, Q. 2010. Product demand forecasts using wavelet kernel support vector machine and particle swarm optimization in manufacture system. *Journal of Computational and Applied Mathematics*, 233(10): 2481–2491.

Xiao, Y. Y., Zhang, R. Q., and Kaku, I. 2011. A new approach of inventory classification based on loss profit. *Expert Systems with Applications*, 38(8): 9382–9391.

Xiong, T., Bao, Y., and Hu, Z. 2014. Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting. *Knowledge-Based Systems*, 55: 87–100.

Yoon, K. P., and Hwang, C. L. 1995. Multiple attribute decision making: an introduction. Sage publications.

Yu, X., Qi, Z., and Zhao, Y. 2013. Support vector regression for newspaper/magazine sales forecasting. *Procedia Computer Science*, 17: 1055–1062.

Zhao, J., and Itti, L. 2018. shapeDTW: Shape dynamic time warping. *Pattern Recognition*, 74(c): 171–184.

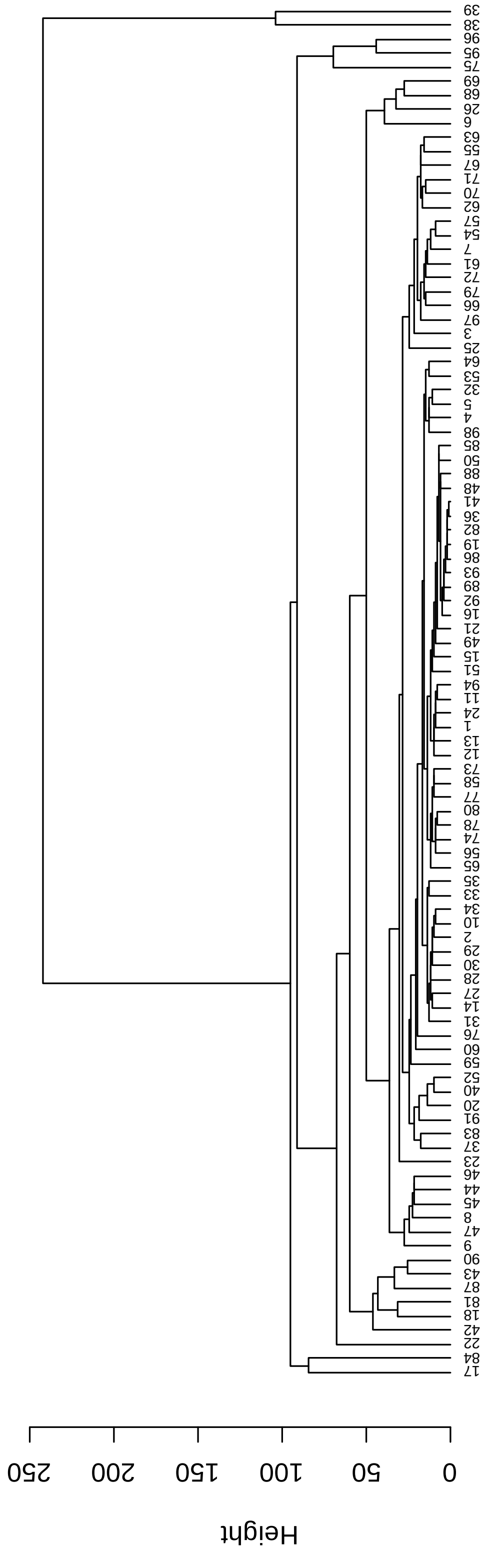
Zhao, X., Xie, J., and Lau, R. S. M. 2001. Improving the supply chain performance: Use of forecasting models versus early order commitments. *International Journal of Production Research*, 39(17): 3923–3939.

Zhou, M., and Wong, M. H. 2005. A segment-wise time warping method for time scaling searching. *Information Sciences*, 173(1–3): 227–254.

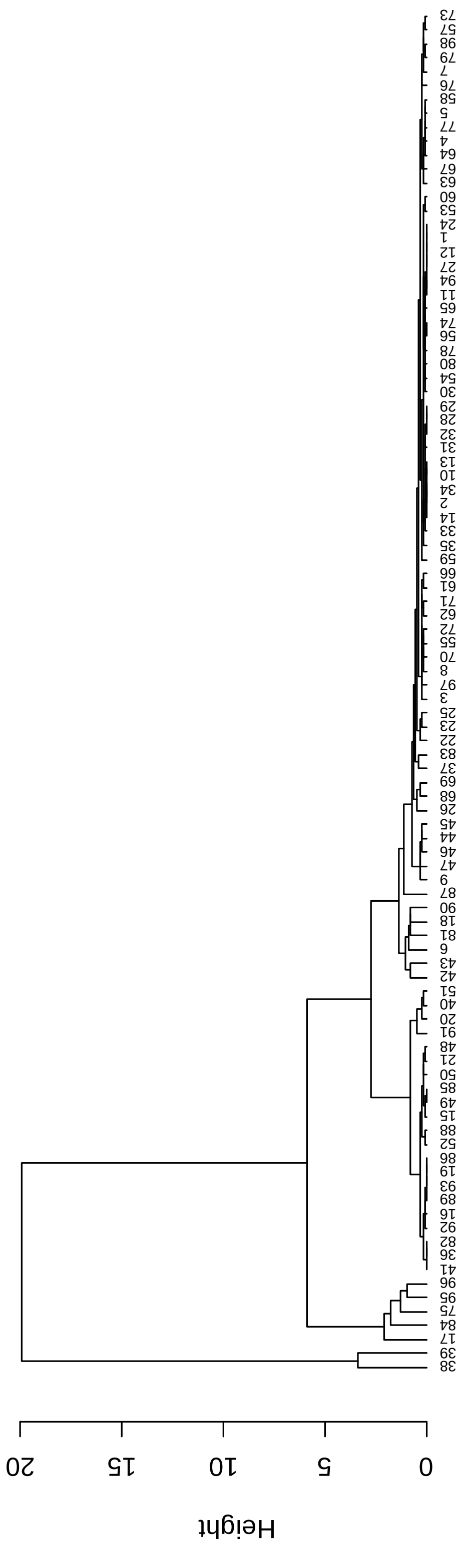
Zuo, Y., Ali, A. B. M. S., and Yada, K. 2014. Consumer purchasing behavior extraction using statistical learning theory. *Procedia Computer Science*, 35: 1464–1473.

APPENDICES

Appendix 1. Dendrogram Using Euclidean Distance



Appendix 2. Dendrogram Using DTW Distance



Appendix 3. Cluster Assignments for *k*=7, *k*= 16, and *k*=27

Item	Cluster		
	k=7	*k*=16	*k*=27
1	1	1	1
2	1	1	1
3	1	1	2
4	1	1	1
5	1	1	1
6	1	2	3
7	1	1	1
8	1	1	2
9	1	1	4
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	2	3	5
16	2	3	5
17	3	4	6
18	1	5	7
19	2	3	5
20	2	3	8
21	2	3	5
22	1	1	9
23	1	1	10
24	1	1	1
25	1	1	10
26	1	1	11
27	1	1	1
28	1	1	1
29	1	1	1
30	1	1	1
31	1	1	1
32	1	1	1
33	1	1	1
34	1	1	1
35	1	1	1
36	2	3	5
37	1	1	12

Item	Cluster		
	k=7	*k*=16	*k*=27
38	4	6	13
39	5	7	14
40	2	3	8
41	2	3	5
42	1	8	15
43	1	9	16
44	1	1	17
45	1	1	17
46	1	1	17
47	1	1	17
48	2	3	5
49	2	3	5
50	2	3	5
51	2	3	8
52	2	3	5
53	1	1	1
54	1	1	1
55	1	1	2
56	1	1	1
57	1	1	1
58	1	1	1
59	1	1	1
60	1	1	1
61	1	1	2
62	1	1	2
63	1	1	1
64	1	1	1
65	1	1	1
66	1	1	2
67	1	1	1
68	1	1	18
69	1	1	18
70	1	1	2
71	1	1	2
72	1	1	2
73	1	1	1
74	1	1	1

Appendix 4. Features without MARS

Methods	Prototype	# of models	without MARS	
			#of selected features	Output variable~Input variables
S-SVR	Medoid	1	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		2	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		3	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		4	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		5	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		6	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		8	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		9	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		10	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		11	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		12	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		13	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		14	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		15	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		16	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		17	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		18	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		19	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		20	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		21	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		22	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		23	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		24	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		25	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		26	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		27	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		28	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		29	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM

Methods	Prototype	# of models	without MARS	
			#of selected features	Output variable~Input variables
S-SVR (*Cont.*)	Medoid (*Cont.*)	30	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		31	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		32	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		33	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		34	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		35	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		36	16	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2
		37	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		38	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		39	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		40	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		41	16	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2
		42	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		43	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		44	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		45	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		46	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		47	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		48	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		49	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		50	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		51	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		52	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		53	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		54	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		55	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		56	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		57	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		58	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM

Methods	Prototype	# of models	without MARS	
			#of selected features	Output variable~Input variables
S-SVR (*Cont.*)	Medoid (*Cont.*)	59	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		60	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		61	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		62	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		63	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		64	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		65	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		66	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		67	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		68	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		69	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		70	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		71	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		72	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		73	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		74	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		75	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		76	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		77	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		78	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		79	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		80	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		81	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		82	16	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2
		83	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		84	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		85	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		86	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		87	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM

Methods	Prototype	# of models	without MARS	
			#of selected features	Output variable~Input variables
S-SVR (*Cont.*)	Medoid (*Cont.*)	88	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		89	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		90	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		91	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		92	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		93	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		94	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		95	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		96	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		97	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		98	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
C-SVR with *k*=7	Medoid	1	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		2	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		3	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		4	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		5	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		6	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
C-SVR with *k*=7	DBA	1	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		2	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		3	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		4	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		5	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		6	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		1	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		2	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		3	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		4	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM

Methods	Prototype	# of models	without MARS	
			#of selected features	Output variable~Input variables
C-SVR with k=16	Medoid	5	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		6	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		7	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		8	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		9	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		10	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		11	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		12	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		13	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		14	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		15	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		16	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
C-SVR with k=16	DBA	1	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		2	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		3	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		4	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		5	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		6	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		7	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		8	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		9	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		10	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		11	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		12	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		13	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		14	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		15	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		16	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		1	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM

Methods	Prototype	# of models	without MARS	
			#of selected features	Output variable~Input variables
C-SVR with k=27	Medoid	2	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		3	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		4	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		5	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		6	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		8	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		9	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		10	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		11	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		12	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		13	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		14	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		15	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		16	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		17	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		18	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		19	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		20	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		21	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		22	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		23	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		24	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		25	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		26	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		27	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		1	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		2	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		3	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM

Methods	Prototype	# of models	without MARS	
			#of selected features	Output variable~Input variables
C-SVR with k=27	DBA	4	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		5	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		6	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		8	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		9	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		10	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		11	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		12	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		13	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		14	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		15	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		16	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		17	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		18	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		19	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		20	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		21	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		22	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		23	29	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		24	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		25	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		26	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		27	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM

Appendix 5. Selected Features by MARS for Each Cluster Prototype

Methods	Prototype	# of models	with MARS: #of selected features	with MARS: Output variable~Input variables
S-MARS+SVR	Medoid	1	16	sales~T10+T20+MA3+MA5+MA15+RDP3+BIAS10+BIAS15+ROC10+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		2	10	sales~MA3+MA5+MA10+MA15+RDP3+RDP15+BIAS15+IMS1+IMS2+IMM
		3	17	sales~T2+T3+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMM
		4	10	sales~T2+T15+MA5+MA10+RDP1+ROC15+Disparity5+IML+IMS1+IMM
		5	10	sales~T3+T5+MA10+RDP1+RDP15+ROC5+ROC10+IML+IMS2+IMM
		6	10	sales~T3+T5+T10+MA5+MA10+MA15+RDP1+RDP5+IML+IMM
		7	11	sales~T10+T20+MA3+MA5+RDP3+ROC5+ROC10+ROC15+Disparity10+OSCP5+IMS1
		8	10	sales~T1+T3+T20+MA2+MA5+MA10+RDP10+BIAS10+ROC5+IMM
		9	10	sales~T3+T5+T20+MA10+RDP3+RDP10+BIAS10+ROC15+OSCP5+IMM
		10	14	sales~T3+MA3+MA10+MA15+RDP3+RDP5+BIAS5+BIAS10+BIAS15+OSCP5+IML+IMS1+IMS2+IMM
		11	11	sales~T15+MA5+MA10+MA15+RDP10+ROC15+OSCP5+IML+IMS1+IMS2+IMM
		12	19	sales~T3+T10+T20+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP15+BIAS5+BIAS10+BIAS15+ROC15+Disparity10+OSCP5+IML+IMS2+IMM
		13	17	sales~T5+MA3+MA10+MA15+RDP1+RDP3+RDP5+RDP10+BIAS10+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		14	10	sales~T5+T10+MA2+MA5+MA10+MA15+ROC10+ROC15+IML+IMM
		15	6	sales~T15+MA3+MA15+RDP3+ROC10+ROC15
		16	5	sales~T10+MA5+MA10+BIAS15+IML
		17	5	sales~MA3+MA10+RDP3+ROC5+IML
		18	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		19	17	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2+IMM
		20	5	sales~T10+MA3+IML+IMS2+IMM
		21	4	sales~MA10+MA15+RDP3+BIAS5
		22	20	sales~T2+T3+T5+T10+T15+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+BIAS10+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		23	18	sales~T5+T10+T15+T20+MA5+MA10+MA15+RDP3+BIAS10+ROC5+ROC10+ROC15+Disparity5+OSCP5+IML+IMS1+IMS2+IMM
		24	11	sales~T3+T10+MA2+MA10+RDP5+Disparity5+OSCP5+IML+IMS1+IMS2+IMM
		25	18	sales~T3+T5+T10+T15+T20+MA2+MA5+MA10+MA15+RDP3+RDP5+RDP10+RDP15+Disparity10+OSCP5+IML+IMS2+IMM
		26	10	sales~RDP1+RDP3+RDP10+BIAS10+BIAS15+ROC5+Disparity5+OSCP5+IML+IMM
		27	15	sales~T5+T10+MA3+MA10+RDP1+RDP3+BIAS10+BIAS15+ROC5+ROC15+Disparity5+OSCP5+IML+IMS2+IMM
		28	10	sales~T1+T5+MA5+MA10+MA15+RDP3+ROC15+Disparity10+OSCP5+IMM
		29	10	sales~T3+T5+MA5+MA15+RDP5+RDP10+RDP15+OSCP5+IML+IMM

Methods	Prototype	# of models	with MARS	
			#of selected features	Output variable~Input variables
S-MARS+SVR (*Cont.*)	Medoid (*Cont.*)	30	10	sales~T3+T5+T20+MA3+MA15+RDP15+ROC5+Disparity5+IML+IMM
		31	10	sales~T2+T3+MA10+MA15+RDP5+RDP10+BIAS15+Disparity5+IMS2+IMM
		32	17	sales~T5+T15+T20+MA3+MA5+MA15+RDP3+RDP5+RDP15+BIAS5+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMM
		33	10	sales~T3+T15+T20+MA15+RDP5+BIAS15+ROC15+Disparity10+IML+IMM
		34	17	sales~T20+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP10+BIAS5+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		35	10	sales~MA2+MA3+MA5+MA10+RDP3+RDP5+RDP15+BIAS15+IMS1+IMM
		36	16	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2
		37	6	sales~T10+RDP10+RDP15+ROC5+Disparity5+Disparity10
		38	10	sales~T1+T15+MA2+MA5+MA15+RDP3+RDP5+BIAS15+ROC10+Disparity10
		39	10	sales~T5+T15+T20+MA2+RDP1+RDP3+RDP10+BIAS15+ROC5+Disparity10
		40	4	sales~MA15+RDP3+ROC5+ROC15
		41	16	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2
		42	2	sales~T10+OSCP5
		43	2	sales~T10+ROC10
		44	10	sales~T3+T5+T10+T15+T20+MA2+MA5+RDP10+BIAS10+ROC15
		45	14	sales~T1+T3+T5+MA2+MA15+RDP1+RDP3+RDP10+BIAS15+ROC10+ROC15+Disparity5+Disparity10+IMM
		46	10	sales~T1+T3+MA2+MA5+MA15+RDP3+BIAS10+ROC5+ROC10+IMM
		47	11	sales~T5+T10+T20+MA5+MA10+MA15+RDP3+ROC5+Disparity10+IML+IMM
		48	6	sales~MA3+MA10+RDP3+RDP10+BIAS10+IMM
		49	8	sales~T20+MA10+RDP3+RDP15+BIAS5+IML+IMS2+IMM
		50	4	sales~MA3+RDP1+ROC5+IML
		51	3	sales~T2+IML+IMS2
		52	10	sales~T2+T3+T15+MA5+RDP3+Disparity5+OSCP5+IML+IMS2+IMM
		53	10	sales~T3+T5+T10+MA3+MA5+MA10+RDP3+ROC10+OSCP5+IML
		54	10	sales~T3+T5+T20+MA3+MA5+ROC5+ROC10+ROC15+Disparity5+IML
		55	19	sales~T3+T5+T10+T15+T20+MA3+MA5+MA10+MA15+RDP3+RDP5+RDP15+BIAS5+BIAS10+BIAS15+ROC10+ROC15+OSCP5+IML
		56	10	sales~T2+T20+MA3+MA15+RDP3+RDP10+Disparity5+OSCP5+IML+IMM
		57	10	sales~T3+T5+T10+T15+T20+MA2+RDP3+Disparity5+OSCP5+IMM
		58	18	sales~T5+T10+T15+MA2+MA3+MA5+MA10+MA15+RDP1+BIAS5+BIAS15+ROC10+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM

Methods	Prototype	# of models	with MARS	
			#of selected features	Output variable~Input variables
S-MARS+SVR (*Cont.*)	Medoid (*Cont.*)	59	16	sales~T10+T15+T20+MA3+MA5+MA15+RDP1+RDP3+RDP10+RDP15+BIAS5+BIAS15+OSCP5+IML+IMS2+IMM
		60	20	sales~T5+T15+MA3+MA5+MA10+MA15+RDP1+RDP3+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+OSCP5+IML+IMS1+IMS2+IMM
		61	10	sales~T15+T20+MA3+MA5+MA15+RDP1+RDP10+Disparity10+IML+IMM
		62	10	sales~T10+T15+T20+MA10+MA15+RDP3+BIAS5+ROC10+IML+IMM
		63	19	sales~T3+T5+T10+T15+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS10+ROC5+Disparity5+OSCP5+IML+IMM
		64	10	sales~T3+T15+MA3+MA10+MA15+RDP3+BIAS10+ROC5+ROC10+IML
		65	12	sales~T2+T3+T10+T20+MA10+MA15+RDP3+ROC10+ROC15+IML+IMS1+IMM
		66	10	sales~T5+MA2+MA3+MA5+RDP1+RDP3+BIAS15+Disparity10+IML+IMM
		67	10	sales~T10+T15+MA3+MA5+MA10+MA15+RDP1+ROC10+OSCP5+IMM
		68	13	sales~T2+T5+T10+MA2+MA3+MA10+MA15+RDP1+RDP5+BIAS15+Disparity10+IML+IMS2
		69	10	sales~T1+T3+T5+MA3+MA10+MA15+RDP3+RDP15+BIAS5+BIAS15
		70	13	sales~T10+T15+T20+MA3+MA5+MA10+MA15+RDP3+BIAS15+ROC5+Disparity10+IML+IMM
		71	10	sales~T2+T5+T20+MA10+MA15+RDP1+BIAS5+ROC5+OSCP5+IMM
		72	18	sales~T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP3+BIAS5+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMM
		73	12	sales~T3+T5+MA2+MA5+MA10+MA15+RDP1+RDP10+ROC15+IML+IMS2+IMM
		74	10	sales~T3+T10+T20+MA10+RDP3+RDP5+RDP10+BIAS15+Disparity5+IMM
		75	6	sales~T3+T20+MA10+ROC15+OSCP5+IML
		76	11	sales~T2+T3+T10+T15+MA15+RDP3+BIAS10+BIAS15+Disparity5+IMS1+IMS2
		77	19	sales~T5+MA3+MA5+MA10+MA15+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+Disparity5+OSCP5+IML+IMS2+IMM
		78	10	sales~T3+T5+T20+MA3+MA10+MA15+ROC5+ROC15+IML+IMM
		79	22	sales~T3+T5+T10+T20+MA2+MA3+MA5+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS15+ROC5+ROC10+ROC15+Disparity5+OSCP5+IML+IMS2+IMM
		80	10	sales~T2+T3+MA5+MA10+BIAS15+ROC5+Disparity5+Disparity10+IML+IMM
		81	2	sales~T5+RDP15
		82	16	sales~T1+T2+T3+T5+MA2+MA3+MA5+RDP1+RDP3+RDP5+BIAS5+ROC5+Disparity5+IML+IMS1+IMS2
		83	5	sales~T5+ROC5+ROC15+Disparity10+OSCP5
		84	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		85	5	sales~T5+RDP5+ROC15+IMS2+IMM
		86	4	sales~T5+RDP3+RDP5+IML
		87	3	sales~T10+MA10+MA15

Methods	Prototype	# of models	with MARS	
			#of selected features	Output variable~Input variables
S- MARS+SVR (*Cont.*)	Medoid (*Cont.*)	88	2	sales~MA3+IML
		89	2	sales~MA5+MA10
		90	3	sales~T20+RDP1+IML
		91	2	sales~RDP1+RDP3
		92	2	sales~MA3+IMS2
		93	3	sales~RDP5+IMS1+IMS2
		94	14	sales~T15+MA3+MA10+MA15+RDP1+RDP3+BIAS5+BIAS15+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		95	4	sales~T15+RDP3+BIAS15+OSCP5
		96	5	sales~T5+T20+RDP10+IML+IMM
		97	19	sales~T2+T3+T10+T15+T20+MA5+MA10+MA15+RDP3+RDP5+RDP10+BIAS5+BIAS10+BIAS15+ROC10+Disparity10+IML+IMS2+IMM
		98	18	sales~T2+T3+T5+T10+T15+MA3+MA5+MA10+MA15+RDP3+RDP5+RDP10+BIAS5+BIAS10+BIAS15+OSCP5+IML+IMM
C- MARS+SVR with *k*=7	Medoid	1	19	sales~T2+T3+T10+T15+T20+MA5+MA10+MA15+RDP3+RDP5+RDP10+BIAS5+BIAS10+BIAS15+ROC10+Disparity10+IML+IMS2+IMM
		2	2	sales~MA3+IMS2
		3	5	sales~MA3+MA10+RDP3+ROC5+IML
		4	10	sales~T1+T15+MA2+MA5+MA15+RDP3+RDP5+BIAS15+ROC10+Disparity10
		5	10	sales~T5+T15+T20+MA2+RDP1+RDP3+RDP10+BIAS15+ROC5+Disparity10
		6	4	sales~T15+RDP3+BIAS15+OSCP5
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
C- MARS+SVR with *k*=7	DBA	1	1	sales~MA10
		2	3	sales~IMS2+RDP1+IML
		3	5	sales~MA3+MA10+RDP3+ROC5+IML
		4	10	sales~T1+T15+MA2+MA5+MA15+RDP3+RDP5+BIAS15+ROC10+Disparity10
		5	10	sales~T5+T15+T20+MA2+RDP1+RDP3+RDP10+BIAS15+ROC5+Disparity10
		6	10	sales~ROC5+MA10+MA15+MA3+T10+T15+T1+BIAS15+ROC15+ROC10
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		1	19	sales~T2+T3+T10+T15+T20+MA5+MA10+MA15+RDP3+RDP5+RDP10+BIAS5+BIAS10+BIAS15+ROC10+Disparity10+IML+IMS2+IMM
		2	10	sales~T3+T5+T10+MA5+MA10+MA15+RDP1+RDP5+IML+IMM
		3	2	sales~MA3+IMS2
		4	5	sales~MA3+MA10+RDP3+ROC5+IML

Methods	Prototype	# of models	with MARS	
			#of selected features	Output variable~Input variables
C-MARS+SVR with *k*=16	Medoid	5	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		6	10	sales~T1+T15+MA2+MA5+MA15+RDP3+RDP5+BIAS15+ROC10+Disparity10
		7	10	sales~T5+T15+T20+MA2+RDP1+RDP3+RDP10+BIAS15+ROC5+Disparity10
		8	2	sales~T10+OSCP5
		9	2	sales~T10+ROC10
		10	6	sales~T3+T20+MA10+ROC15+OSCP5+IML
		11	2	sales~T5+RDP15
		12	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		13	3	sales~T10+MA10+MA15
		14	3	sales~T20+RDP1+IML
		15	4	sales~T15+RDP3+BIAS15+OSCP5
		16	5	sales~T5+T20+RDP10+IML+IMM
C- MARS+SVR with *k*=16	DBA	1	7	sales~IMS1+RDP5+IMS2+T5+OSCP5+IMM+IML
		2	10	sales~T3+T5+T10+MA5+MA10+MA15+RDP1+RDP5+IML+IMM
		3	3	sales~IMS2+RDP1+IML
		4	5	sales~MA3+MA10+RDP3+ROC5+IML
		5	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		6	10	sales~T1+T15+MA2+MA5+MA15+RDP3+RDP5+BIAS15+ROC10+Disparity10
		7	10	sales~T5+T15+T20+MA2+RDP1+RDP3+RDP10+BIAS15+ROC5+Disparity10
		8	2	sales~T10+OSCP5
		9	2	sales~T10+ROC10
		10	6	sales~T3+T20+MA10+ROC15+OSCP5+IML
		11	2	sales~T5+RDP15
		12	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		13	3	sales~T10+MA10+MA15
		14	3	sales~T20+RDP1+IML
		15	4	sales~T15+RDP3+BIAS15+OSCP5
		16	5	sales~T5+T20+RDP10+IML+IMM
		1	14	sales~T15+MA3+MA10+MA15+RDP1+RDP3+BIAS5+BIAS15+Disparity10+OSCP5+IML+IMS1+IMS2+IMM

Methods	Prototype	# of models	with MARS	
			#of selected features	Output variable~Input variables
C-MARS+SVR with k=27	Medoid	2	18	sales~T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP3+BIAS5+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMM
		3	10	sales~T3+T5+T10+MA5+MA10+MA15+RDP1+RDP5+IML+IMM
		4	10	sales~T3+T5+T20+MA10+RDP3+RDP10+BIAS10+ROC15+OSCP5+IMM
		5	2	sales~MA3+IMS2
		6	5	sales~MA3+MA10+RDP3+ROC5+IML
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		8	4	sales~MA15+RDP3+ROC5+ROC15
		9	20	sales~T2+T3+T5+T10+T15+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+BIAS10+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		10	18	sales~T5+T10+T15+T20+MA5+MA10+MA15+RDP3+BIAS10+ROC5+ROC10+ROC15+Disparity5+OSCP5+IML+IMS1+IMS2+IMM
		11	10	sales~RDP1+RDP3+RDP10+BIAS10+BIAS15+ROC5+Disparity5+OSCP5+IML+IMM
		12	6	sales~T10+RDP10+RDP15+ROC5+Disparity5+Disparity10
		13	10	sales~T1+T15+MA2+MA5+MA15+RDP3+RDP5+BIAS15+ROC10+Disparity10
		14	10	sales~T5+T15+T20+MA2+RDP1+RDP3+RDP10+BIAS15+ROC5+Disparity10
		15	2	sales~T10+OSCP5
		16	2	sales~T10+ROC10
		17	10	sales~T1+T3+MA2+MA5+MA15+RDP3+BIAS10+ROC5+ROC10+IMM
		18	13	sales~T2+T5+T10+MA2+MA3+MA10+MA15+RDP1+RDP5+BIAS15+Disparity10+IML+IMS2
		19	6	sales~T3+T20+MA10+ROC15+OSCP5+IML
		20	2	sales~T5+RDP15
		21	5	sales~T5+ROC5+ROC15+Disparity10+OSCP5
		22	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		23	3	sales~T10+MA10+MA15
		24	3	sales~T20+RDP1+IML
		25	2	sales~RDP1+RDP3
		26	4	sales~T15+RDP3+BIAS15+OSCP5
		27	5	sales~T5+T20+RDP10+IML+IMM
		1	10	sales~IMM+Disparity5+IMS1+MA3+IMS2+T3+T5+RDP3+RDP1+T2
		2	10	sales~T15+RDP3+RDP15+ROC5+IML+IMM+T5+MA15+BIAS15+RDP1
		3	10	sales~T3+T5+T10+MA5+MA10+MA15+RDP1+RDP5+IML+IMM

Methods	Prototype	# of models	with MARS	
			#of selected features	Output variable~Input variables
C-MARS+SVR with k=27	DBA	4	10	sales~T3+T5+T20+MA10+RDP3+RDP10+BIAS10+ROC15+OSCP5+IMM
		5	3	sales~IMM+RDP1+MA5
		6	5	sales~MA3+MA10+RDP3+ROC5+IML
		7	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		8	4	sales~T10+IML+MA3+MA15
		9	20	sales~T2+T3+T5+T10+T15+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+BIAS10+BIAS15+Disparity5+Disparity10+OSCP5+IML+IMS2+IMM
		10	14	sales~IML+MA10+MA15+T15+MA5+OSCP5+T10+IMM+RDP5+BIAS5+BIAS10+BIAS15+RDP15+MA2
		11	10	sales~RDP1+RDP3+RDP10+BIAS10+BIAS15+ROC5+Disparity5+OSCP5+IML+IMM
		12	6	sales~T10+RDP10+RDP15+ROC5+Disparity5+Disparity10
		13	10	sales~T1+T15+MA2+MA5+MA15+RDP3+RDP5+BIAS15+ROC10+Disparity10
		14	10	sales~T5+T15+T20+MA2+RDP1+RDP3+RDP10+BIAS15+ROC5+Disparity10
		15	2	sales~T10+OSCP5
		16	2	sales~T10+ROC10
		17	15	sales~IML+T3+MA15+BIAS15+RDP3+IMS1+T10+ROC5+IMS2+T2+IMM+ROC15+ROC10+T5+MA2
		18	10	sales~T15+OSCP5+IML+MA15+T2+T20+MA10+MA5+T10+IMM
		19	6	sales~T3+T20+MA10+ROC15+OSCP5+IML
		20	2	sales~T5+RDP15
		21	5	sales~T5+ROC5+ROC15+Disparity10+OSCP5
		22	30	sales~T1+T2+T3+T5+T10+T15+T20+MA2+MA3+MA5+MA10+MA15+RDP1+RDP3+RDP5+RDP10+RDP15+BIAS5+BIAS10+BIAS15+ROC5+ROC10+ROC15+Disparity5+Disparity10+OSCP5+IML+IMS1+IMS2+IMM
		23	3	sales~T10+MA10+MA15
		24	3	sales~T20+RDP1+IML
		25	2	sales~RDP1+RDP3
		26	4	sales~T15+RDP3+BIAS15+OSCP5
		27	5	sales~T5+T20+RDP10+IML+IMM

Appendix 6. The Rules Generated by the Decision Tree

Number of clusters (***k***)	Cluster No.	Coverage (***n***)	Rule Number*	Description
7	1	7634	1	IML < 0.7795
				IML >= 0.133
	2	1178	1	IML >= 0.7795
	4	46	1	IML < 0.7795
				IML < 0.133
				IML < 0.04587
				MA5 < 22.3
	5	82	1	IML < 0.7795
				IML < 0.133
				IML < 0.04587
				MA5 >= 22.3
	6	28	1	IML < 0.7795
				IML < 0.133
				IML >= 0.04587
16	1	7648	1	IML < 0.7864
				IML >= 0.133
	3	1164	1	IML >= 0.7864
	6	46	1	IML < 0.7864
				IML < 0.133
				IML < 0.04587
				MA5 < 22.3
	7	82	1	IML < 0.7864
				IML < 0.133
				IML < 0.04587
				MA5 >= 22.3
	15	28	1	IML < 0.7864
				IML < 0.133
				IML >= 0.04587
27	1	912	1	IML < 0.7963
				IML < 0.4069
				IML >= 0.323
				MA15 >= 1.367
	1	5432	2	IML < 0.7963
				IML >= 0.4069
	2	803	1	IML < 0.7963
				IML < 0.4069
				IML < 0.323
				IML >= 0.04587
				MA15 < 2.833

Number of clusters (*k*)	**Cluster No.**	**Coverage (*n*)**	**Rule Number***	**Description**
27 (*Cont*.)	2	305	2	IML < 0.7963
				IML < 0.4069
				IML >= 0.323
				MA15 < 1.367
	5	1139	1	IML >= 0.7963
	13	128	1	IML < 0.7963
				IML < 0.4069
				IML < 0.323
				IML < 0.04587
	17	249	1	IML < 0.7963
				IML < 0.4069
				IML < 0.323
				IML >= 0.04587
				MA15 >= 2.833

*some clusters may have multiple rules

Appendix 7. Minitab Outputs of Wilcoxon Signed Rank Test

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=27 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	18,8475

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	97	4753,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	21,8733

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	97	4753,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and medoid (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	27,4036

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	97	4753,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and medoid (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	31,2057

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	98	4851,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	29,4770

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	98	4851,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	18,0807

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	96	4656,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and medoid (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	21,3687

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	97	4753,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and medoid (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	21,9608

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	98	4851,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	19,9352

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	98	4851,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=16 and medoid (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	18,1996

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	84	3570,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and medoid (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	21,8576

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	98	4851,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	21,1167

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	98	4851,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and medoid (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	16,2593

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	97	4753,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	16,9299

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	98	4851,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=7 and medoid vs. C-MARS+SVR for *k*=7 and DBA (MAPE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	98	17,8541

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	94	4465,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=27 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,280275

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	76	2926,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,456047

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	83	3486,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and medoid (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,341493

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	83	3486,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and medoid (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,412396

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	92	4278,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,613562

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	93	4371,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,169130

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	82	3403,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and medoid (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,105884

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	82	3403,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and medoid (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,166315

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	91	4186,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,264685

Test

Null hypothesis H$_0$: η = 0

Alternative hypothesis H$_1$: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	92	4278,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=16 and medoid (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,166933

Test

Null hypothesis H$_0$: η = 0

Alternative hypothesis H$_1$: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	82	3403,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and medoid (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,216318

Test

Null hypothesis H$_0$: η = 0

Alternative hypothesis H$_1$: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	91	4186,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,0258874

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	69	2415,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and medoid (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	9	45,00	0,009

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,225959

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	90	4095,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=7 and medoid vs. C-MARS+SVR for *k*=7 and DBA (RMSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,212623

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	90	4095,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=27 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,388889

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	76	2926,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,764768

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	83	3486,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and medoid (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,495686

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	83	3486,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and medoid (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,604514

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	92	4278,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	1,10952

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	93	4371,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,389277

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	82	3403,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and medoid (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,207353

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	82	3403,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and medoid (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,297619

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	91	4186,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,600917

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	92	4278,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=16 and medoid (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,366569

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	82	3403,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and medoid (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,533284

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	91	4186,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,056

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	69	2415,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and medoid (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	9	45,00	0,009

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,571429

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	90	4095,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=7 and medoid vs. C-MARS+SVR for *k*=7 and DBA (MSE)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,526357

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	90	4095,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=27 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,247573

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	76	2926,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,370833

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	83	3486,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=16 and medoid (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,323672

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	83	3486,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and medoid (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,401623

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	92	4278,00	0,000

Wilcoxon Signed Rank Test:
S-MARS+SVR vs. C-MARS+SVR for *k*=7 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,524075

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	93	4371,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,110125

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	78	3081,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=16 and medoid (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,145215

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	82	3403,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and medoid (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,214286

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	91	4186,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=27 and DBA vs. C-MARS+SVR for *k*=7 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,237554

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	91	4186,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=16 and medoid (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,106157

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	80	3240,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and medoid (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,150156

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	89	4005,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and DBA vs. C-MARS+SVR for *k*=7 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,054923

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	69	2415,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and medoid (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	9	45,00	0,009

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=16 and medoid vs. C-MARS+SVR for *k*=7 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,183333

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	92	4278,00	0,000

Wilcoxon Signed Rank Test:
C-MARS+SVR for *k*=7 and medoid vs. C-MARS+SVR for *k*=7 and DBA (MAD)

Method

η: median of Difference

Descriptive Statistics

Sample	N	Median
Difference	97	0,170250

Test

Null hypothesis	H_0: η = 0
Alternative hypothesis	H_1: η ≠ 0

Sample	N for Test	Wilcoxon Statistic	P-Value
Difference	91	4186,00	0,000

Appendix 8. Evaluation of Inventory Performance for Item 26

item **26**
inventory policy **LFL**

Initial inventory	2	units	holding factor	11.0%	per unit/year
Lead time (LT)	3	month	Purchased price	10840	$/unit
Safety stock (SS)	**0**	**unit**	Order cost	13500	$/order
Total cost	131117.00	$	Holding cost	99	$/unit/month
Total backorder	7		Backorder cost	434	$/unit/month
Total order place	9				

Month	Begining inventory	Receive	Actual demand	Backorder unit	Accumulated backorder unit	Ending inventory	Inventory on-order	Forecasted demand (t+1)	Forecasted demand (t+2)	Forecasted demand (t+3)	Forecasted demand (t+4)	Forecasted demand (t+5)	Forecasted demand (t+6)	Projected Inventory Position	Order placement	Order quantity	Holding cost	Order cost	Backorder cost	Total cost
1	2	0	0	0	0	2	0	2	2	2	2	2	2	-4	YES	6.00	$198	$13,500	$0	$13,698
2	2	0	2	0	0	0	6	2	1	1	1	1	1	2	NO	0.00	$0	$0	$0	$0
3	0	0	1	1	1	0	6	2	1	1	1	1	1	1	NO	0.00	$0	$0	$434	$434
4	0	0	1	1	2	0	0	2	1	1	1	1	1	-6	YES	7.00	$0	$13,500	$434	$13,934
5	0	6	6	0	2	0	7	2	1	1	1	1	1	1	NO	0.00	$0	$0	$0	$0
6	0	0	3	3	5	0	7	2	1	1	1	1	1	-2	YES	3.00	$0	$13,500	$1,302	$14,802
7	0	0	1	1	6	0	3	2	1	1	1	1	1	-7	YES	8.00	$0	$13,500	$434	$13,934
8	0	7	1	0	0	0	11	2	1	1	1	1	1	7	NO	0.00	$0	$0	$0	$0
9	0	0	1	1	1	0	8	2	1	1	1	1	1	3	NO	0.00	$0	$0	$434	$434
10	0	3	0	0	0	2	0	2	1	1	1	1	1	-2	YES	3.00	$198	$13,500	$0	$13,698
11	2	8	2	0	0	8	3	2	1	1	1	1	1	7	NO	0.00	$792	$0	$0	$792
12	8	0	13	5	5	0	3	2	1	1	1	1	1	-6	YES	7.00	$0	$13,500	$2,170	$15,670
13	0	0	4	4	9	0	7	2	1	1	1	1	1	-6	YES	7.00	$0	$13,500	$1,736	$15,236
14	0	3	1	0	7	0	14	2	1	1	1	1	1	3	NO	0.00	$0	$0	$0	$0
15	0	0	0	0	7	0	7	2	1	1	1	1	1	-4	YES	5.00	$0	$13,500	$0	$13,500
16	0	7	0	0	0	0	5	2	1	1	1	1	1	1	NO	0.00	$0	$0	$0	$0
17	0	7	0	0	0	7	5	2	1	1	1	1	1	8	NO	0.00	$693	$0	$0	$693
18	7	0	3	0	0	4	0	2	2	1	1	1	1	-1	YES	2.00	$396	$13,500	$0	$13,896
19	4	5	5	0	0	4	2	2	2	1	1	1	1	1	NO	0.00	$396	$0	$0	$396

CURRICULUM VITAE

Name, Surname	: Pratiwi Eka Puspita
Date and Place of Birth	: Temanggung, 22nd February 1989
Foreign Language	: English, Turkish

Education Background (Institution and Year)

High School	: SMA N 1 Temanggung, 2006
Bachelor of Science	: Bogor Agricultural University, 2011

Working Experience / Institution and Year	: -
Email	: pratiwiekapuspita@gmail.com
Publications	:

Puspita, P. E., İnkaya, T., and Akansel, M. 2017. Clustering-based sales forecasting in a forklift distributor. Yöneylem Araştırması ve Endüstri Mühendisliği Ulusal Kongresi, 5-7 July 2017, Ankara.