Journal of Optimization in Industrial Engineering
Vol.13, Issue 2, Summer & Autumn 2020, 101-121
DOI: 10.22094/JOIE.2020.1863364.1630

Research Article

A Data Mining Approach for Forecasting Failure Root Causes: A Case Study in an Automated Teller Machine (ATM) Manufacturing Company

Seyedeh Pardis Bagherighadikolaei[a], Rouzbeh Ghousi [a,*], Abdorrahman Haeri[a]

[a] *Department of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran*
Received 25 February 2019; Revised 25 January2020; Accepted 23 February 2020

Abstract

Based on the findings of Massachusetts Institute of Technology, organizations' data double every five years. However, the rate of using data is 0.3. Nowadays, data mining tools have greatly facilitated the process of knowledge extraction from a welter of data. This paper presents a hybrid model using data gathered from an ATM manufacturing company. The steps of the research are based on CRISP-DM. Therefore, based on the first step, business understanding, the company and its different units were studied. After business understanding, the data collected from sale's unit were prepared for preprocess. While preprocessing, data from some columns of dataset, based on their types and purpose of the research, were either categorized or coded. Then, the data have been inserted into Clementine software, which resulted in modeling and pattern discovery. The results clearly state that, the same Machines' Code and the same customers in different provinces are struggling with significantly different Problems' Code, that could be due to weather condition, culture of using ATMs, and likewise. Moreover, the same Machines' Code and the same Problems' Code, as well as differences in Technicians' expertise, seems to be some causes to significantly different Repair Time. This could be due to Technicians' training background level of their expertise and such. At last, the company can benefit from the outputs of this model in terms of its strategic decision-making.

Keywords: Data Mining; Clustering; Association Rules; classification; Automated Teller Machine (ATM).

1. Introduction

By the advent of computers capable of restoring and analyzing data, the technology has been developed, resulting in gradual increase of the amount of restored data. Internet, integrated information systems, and E-commerce play an important role in adding data to databases. As a result, the databases has turned into huge data sets. Therefore, it is necessary to precisely discover and extract knowledge from the large databases. No need to say, the serious competitions in scientific, social, economic, political and military fields have a profound effect on the speed of accessibility to data. Hence, one can clearly perceive the growing need to design systems, being able to access quickly to the users' required information, minimizing the role of humans in this process, and incorporate appropriate methods, easing the access to large databases. Currently, Data Mining is the most prominent technology for the effective, correct, and quick use of large data sets. The discovery of effective patterns happens in the process of data mining as a valid, simple, and understandable model, describing the relations among subsets of data.
In order to enhance the ATMs' performance and increase customer satisfaction, this article aims to present a hybrid model, using data mining tools such as clustering, association rules, and classification. The discovered relations among data help managers to make more effective and accurate decisions.

1.1. Data Mining

Data Mining deals with various fields such as databases, statistics, as well as machine learning to extract hidden and valuable knowledge and information from a welter of data. Data Mining is one of the methods, through which the effective patterns of data with the least interference of users are discovered. In addition, the outputs of using data mining methods enable managers to make decisively crucial decisions (Han et al., 2011).
Although there are several definitions of data mining, three selective definitions could be found below:
1- The process of extracting knowledge and information and discovering hidden patterns in complicatedly large data sets (Berry & Linoff, 1997).
2- The process of analyzing computer descriptive data in complicatedly large data sets (Friedman et al., 2001).

3- A newly developed interdisciplinary dealing with various fields such as databases, statistics, machine learning to extract hidden and valuable knowledge and information from a welter of data (Ngai et al., 2009).
There are seven Data Mining tools: Association Rules, Classification, Clustering, Prediction, Regression, Sequence Discovery, and Visualization (Ahmed, 2001; Giraud-Carrier & Povel, 2003; Mitra et al., 2002; Shaw et al., 2001).

*Corresponding author Email address: ghousi@iust.ac.ir

2. Literature Review

Despite, growing number of people using ATMs to manage their financial affairs, researches done show ATM maintenance data has not been analyzed thoroughly. When a customer inserts his/her credit card into an ATM, he/she faces a machine not a human. Therefore, it greatly matters to win his/her trust by providing appropriate services, such as an immediate effective performance of an ATM. Considering the importance of ATMs' performance, the culprits of ATMs breakdowns and their correlation must be identified to decrease the repair time and minimize the frequency and intensity of breakdowns. Table 1**Error! Reference source not found.** shows articles from 2008 to 2018, which have been studied in the current article regarding Data Mining tools used in, as well as their scope of application.

Table 1
Articles studied in this work in which Data Mining tools used

Studying articles in which Data Mining tools have been used																			
									Application Field										
Author	**Clustering**	**Association Rules**	**Classification**	**Clustering and Classification**	**Clustering and Association Rules**	**Classification and Association Rules**	**Hybrid of Clustering Algorithms**	**Hybrid of Classification Algorithms**	**Transportation (Automobile, Accident)**	**Healthcare**	**Bank**	**Food Industry**	**Hygienic Industry**	**Gas and Oil Industry**	**Retail**	**Manufacturing**	**Online Shopping Markets**	**Electronic Industry**	**Agriculture**
(Kim & Ahn, 2008)	■																■		
(Mahdavi et al., 2008)				■													■		
(Shih & Liu, 2008)					■										■				
(İşeri & Karlık, 2009)			■						■										
(Cheng & Chen, 2009)				■														■	
(Karabatak & Ince, 2009)						■				■									
(Hosseini et al., 2010)	■								■										
(Mirabadi & Sharifian, 2010)		■							■										
(Tsai & Chen, 2010)						■												■	
(Rumpf et al., 2010)								■											■
(Movagharnejad et al., 2011)			■											■					
(Cimpoiu et al., 2011)			■									■							
(Chian, 2011)					■												■		
(Parvaneh et al., 2012)	■												■						
(Golmah & Mirhashemi, 2012)	■															■			
(Dhandayudam & Krishnamurthi, 2012)	■															■			
Studying articles in which Data Mining tools have been used																			
Author									**Application Field**										

	Clustering	Association Rules	Classification	Clustering and Classification	Clustering and Association Rules	Classification and Association Rules	Hybrid of Clustering Algorithms	Hybrid of Classification Algorithms	Transportation (Automobile, Accident)	Healthcare	Bank	Food Industry	Hygienic Industry	Gas and Oil Industry	Retail	Manufacturing	Online Shopping Markets	Electronic Industry	Agriculture
(Kargari & Sepehri, 2012)	■															■			
(Cil, 2012)		■													■				
(de Oña et al., 2012)			■						■										
(Chen et al., 2012)				■						■									
(Balaji & Srivatsa, 2012)					■					■									
(Shim et al., 2012)						■											■		
(Palomares-Salas et al., 2014)			■																■
(Silvera et al., 2014)			■							■									
(Vafaie et al., 2014)			■							■									
(Venkatesh et al., 2014)				■							■								
(Seret et al., 2014)							■								■				
(Videla-Cavieres & Ríos, 2014)							■								■				
(Bhuvaneswari et al., 2014)								■		■									
(Cremaschi et al., 2015)	■													■					
(Arumawadu et al., 2015)	■																	■	
(Vijayarani & Sudha, 2015)	■									■									
(Chen, 2015)		■										■							
(Ivančević et al., 2015)		■								■									
(Zhang J et al., 2015)			■											■					
(da Silva CET et al., 2015)			■									■							
(Keshavarzi et al., 2015)			■																■
(Kareem & Lawal, 2015)			■						■										
(Elyasigomari et al., 2015)				■						■									
(Vermeulen-Smit et al., 2015)				■						■									
Studying articles in which Data Mining tools have been used																			
Author	**st er**	**ti o**	**ifi ca**	**n d C**	**A**	**n d A**	**st er**	**in ca**	**Application Field**										

									Transportation (Automobile, Accident)	Healthcare	Bank	Food Industry	Hygienic Industry	Gas and Oil Industry	Retail	Manufacturing	Online Shopping Markets	Electronic Industry	Agriculture
(Manolopoulou et al., 2015)						■					■								
(Wang et al., 2016)	■															■			
(Molina et al., 2016)	■									■									
(Weng et al., 2016)		■							■										
(Doub et al., 2016)		■										■							
(Zhao et al., 2016)			■							■									
(Sudha et al., 2016)			■									■							
(Nasibov et al., 2016)			■									■							
(Wang et al., 2017)						■				■									
(Luo et al., 2017)								■		■									
(Martin & Bestle, 2018)			■													■			
(Kuo & Zulvia, 2018)							■								■				
(Lee et al., 2018)								■									■		

3. Materials and Methods

Clustering, Association Rules, and Classification are three Data Mining tools, have been used in this paper.

3.1. K-means Clustering Method

Clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). Additionally, there are K similar groups in K-means algorithm. It is necessary to determine the optimum cluster to carry out a careful analysis. There are various indexes for determining the optimum clusters (Han et al., 2011, Berry & Linoff, 1997, Tsiptsis & Chorianopoulos, 2011).

Four indexes will be explained as follows:

3.1.1. Sum of Squared Errors (SSE)

The distance or similarity between an object and center of the cluster to which the object is assigned can be used to measure how well the object belongs to cluster (Han et al., 2011).
To measure this indicator, first the square of all errors must be calculated. Then, sum of these values must be calculated. In Eq (1) Sum of Squared Errors index is defined.

$$SSE = \sum_{i=1}^{K} \sum_{x \in c_i} dist^2(m_i . x) \qquad (1)$$

x is an element in C_i cluster and m_i is a representative of the cluster C_i.

3.1.2. Davies-Bouldin Index

The lowest value of Davies-Bouldin *inde*x indicates the optimal clustering. This index, Davies-Bouldin is defined in Eq (2).

$$\frac{1}{K}\sum_{i=1}^{K} \max(d_{ij}) \quad ; \quad j = 1. \dots . k \ . \ j \neq 1 \qquad (2)$$

Where $d_{ij} = \frac{S_i + S_j}{d(c_i . c_j)}$. K is the number of clusters, S_i is the average distance of all patterns in cluster i to their cluster centroid and $d(c_i,c_j)$ is the distance between the centroids of clusters I and j. of cluster centers C_i and C_j (Gurrutxaga et al., 2011).

3.1.3. Dunn Index

Let d_{min} denote the smallest distance between two elements from different clusters, and d_{max} the largest distance of two patterns from the same cluster. Then the Dunn index, *D*, is defined in Eq (3).

$$DI = \frac{d_{min}}{d_{max}} \tag{3}$$

Clearly, $D \in [0.\infty]$ with larger values of D indicating better clustering (Tormo et al., 2009).

3.1.4. Silhouette

The Silhouette width is a ratio-type index that is based on Silhouette values for every entity y_i measuring how well y_i fits into the cluster to which it is assigned, by comparing the within-cluster cohesion, based on the distance to all entities in the same cluster, to the cluster separation. Silhouette index is defined in Eq (4).

$$s(y_i) = \frac{b(y_i) - a(y_i)}{max\{a(y_i.b(y_i)\}} \tag{4}$$

Where $a(y_i)$ is the average dissimilarity of $y_i \in S_k$ to all other $y_j \in S_k$, $b(y_i)$ the minimum dissimilarity over all clusters S_l, to which y_i is not assigned, of the average dissimilarities to $y_j \in S_l$, $l \neq k$. Therefore, $-1 \leq s(y_i) \leq 1$. If $s(y_i)$ is around zero, the entity y_i could be assigned to another cluster without making cluster cohesion or separation any worse. A negative $s(y_i)$ suggests that y_i's cluster assignment is damaging to cluster cohesion and separation, whereas an $s(y_i)$ closer to 1 means the opposite. We can then quantify the validity of the whole clustering by the Silhouette index, defined as $\frac{1}{N}\sum_{i\in Y} s(y_i)$ (Van Der Schaaf & Kanse, 2004).

3.2. Association Rules

Let $I = \{I_1.I_2.\ldots.I_m\}$ be an itemset. Let *D*, the task-relevant data, be a set of database transactions where each transaction *T* as a nonempty itemset such that $T \subseteq I$. Each transaction is associated with an identifier, called a *TID*. Let *A* be a set of items. *A* transaction *T* is said to contain *A* if $A \subseteq T$. An association rule is an implication of the form $A \Rightarrow B$, where $A \subset I$, $B \subset I$, $A \neq \phi$, $B \neq \phi$, and $A \cap B = \emptyset$. The rule $A \Rightarrow B$ holds in the transaction set *D* with support *s*, where *s* is the percentage of transactions in *D* that contain $A \cup B$. This is taken to be the probability, $P(A \cup B)$. The rule $A \Rightarrow B$ has confidence *c* in the transaction set *D*, where c is the percentage of transactions in *D* containing *A* that also contain *B*. This is taken to be the conditional probability, $P(B|A)$. Support and confidence indexes are defined in Eq (5) and Eq (6).

$$Support(A \Rightarrow B) = P(A \cup B) \tag{5}$$

$$Confidence(A \Rightarrow B) = P(B|A) = \frac{Support(A \cup B)}{Support(A)} \tag{6}$$

Rules that satisfy both a minimum support threshold *(min_sup)* and a minimum confidence threshold *(minconf)* are called strong.

The support and confidence measures are insufficient at filtering out uninteresting association rules. To tackle this weakness, a correlation measure can be used to augment the support_confidence framework for association rules. This leads to correlation rules of the form; $A \Rightarrow B[support.confidence.correlation]$. That is, a correlation rule is measured not only by its support and confidence, but also by the correlation between items *A* and *B*. Lift is a simple correlation measure that is given as follows. The occurrence of itemset *A* is independent of the occurrence of itemset *B* if $P(A \cup B) = P(A)P(B)$; otherwise, itemsets *A* and *B* are dependent and correlated as events. This definition can easily be extended to more than two itemsets. The lift between the occurrence of *A* and *B* can be measured by computing Eq (7).

$$lift(A.B) = \frac{P(A \cup B)}{P(A)P(B)} \tag{7}$$

If lift index is less than 1, then the occurrence of *A* is negatively correlated with the occurrence of *B*, meaning that the occurrence of one likely leads to the absence of the other one. If the resulting value is greater than 1, then *A* and *B* are positively correlated, meaning that the occurrence of one implies the occurrence of the other. If the resulting value is equal to 1, then *A* and *B* are independent, and there is no correlation between them (Han et al., 2011).

3.3. C5.0 Classification Method

C5.0 is a process, which discovers a model capable of predicting unknown objects by identifying categories or data. Decision Tree plays the most essential piece in C5.0 algorithm and is a common and powerful tool to predict and classify data. Moreover, Decision Tree produces rules within its structure, which explain the obtained results (Han et al., 2011, Berry & Linoff, 1997, Tsiptsis & Chorianopoulos, 2011).

Research Framework, which will be explained base on CRISP-DM method, includes Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Development.

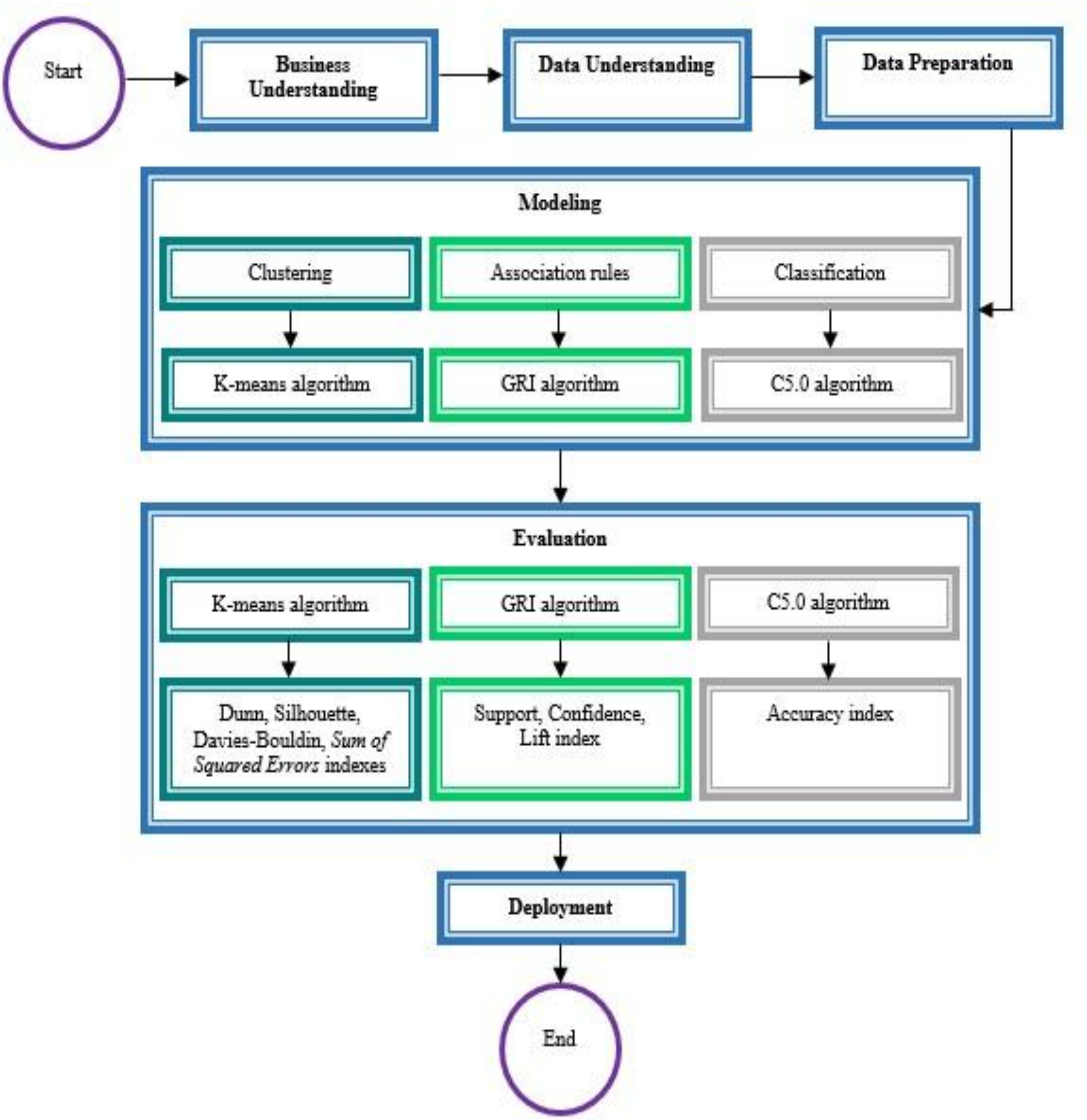

Fig 1. Procedure of research

4. Research Framework

The present article aims to analyze data of ATMs' repairs, in order to offer practical suggestions to company directors, helping them make accurate decisions, leading them to formulate effective strategies. In the initial stages of this article, what the customers mean to accomplish from the business perspective, has been defined via using CRISP-DM methodology. Then, related data of the company (this article's case study) has been studied and preprocessed to prepare them as the software's inputs. The next phase began with Data Modeling. In the process of modeling, K-means algorithm was used for clustering, GRI (Generalized Rule Induction) for discovering rules between data and, C5.0 for classifying data. Then, the evaluation process began. In the next phase, special indexes of each algorithm were used to evaluate the result of each algorithm. We used four indexes, including Dunn Index (DI), Davis-Bouldin Index (DBI), Sum of Squared Error (SSE), and Silhouette index to determine the optimum cluster number, which affects quality of clusters. In addition, using Confidence Index, Support Index and Lift Index, Association Rules were evaluated. Finally, Accuracy Index was used, thereafter, to validate the classification model. Additionally, seven algorithms including: C5, C&R Tree, CHAID, Bayes, Neural Network (NN), Logistic and Support Vector Machine (SVM) were used to classify data, and the accuracy of each algorithm that has been calculated by the software; therefore, the most appropriate algorithm has been chosen. After the evaluation of models and values of each index in the last stage, the model mean to be delivered to the company to be deployed. Fig 1 shows the procedure of this research.

4.1. Business Understanding

The current research is conducted based on an ATMs manufacturing company data. The company imports the pieces of ATMs from Sweden and Korea, assembles the pieces, and sells them all over Iran. The company has technicians, warehouses, and workshops in each province of Iran. The company's customers are largely banks, in case of ATM breakdown, banks would call the company; company's technicians would be send to repair the machines. During the ATM repair service, company's technicians would fill several forms. The current article is going to analyze the related data in this process.

4.2. Data Understanding

In data understanding phase, data collection, codification and quality verification play critical roles in terms of ensuring qualified results in data mining process. An unbiased and accurate data collection leads to more precise analysis results. Different researchers have investigated the causes of errors in data collection and codification stage and proposed models and procedures to limit such dilemmas (Tormo et al. 2009; Van Der Schaaf & Kanse, 2004). In this research, data has been collected in 16863 rows and 45 columns in an excel spreadsheet within two months. Table 2 introduces all of the features.

Table 2.
Features description

Column	**Description**
A	Request Number: Any contact with the company receives a specific request number.
B	Date of Call: The date on which the customer has made contact with the company and made their request.
C	Time of Call: The time at which the customer has made contact with the company.
D	Operator Code: The code of the person who is assigned to answer the customer.
E	Customer Code: Customer code is a number that is dedicated to each customer.
F	Customer Name: It is a column including names of banks, financial institutions and other customers.
G	Code of the Bank's Branch: It is a specific code dedicated to each branches (if the customers have branches).
H	ATM Code: It is determined based on the different models of ATMs.
I	ATM Serial Number: After assembling, the specific serial number is dedicated to each ATM machine.
J	Name of the Province: It is a column showing the province name, in which the customer call has been made.
K	Name of the City: It is a column showing the city name, in which the customer call has been made.
L	ATM Grade: It indicates the rate of transaction for each customer.
M	Log Status: This column cells are filled with two different statements, "normal status (ATM trouble)" and "damaged (caused by people) ".
N	Code of the Trouble: Each trouble has been marked by a code.
O	Problem Statement: Each ATM's trouble is described in this column.
P	ATM Status: This column indicates the details of ATM's guarantee.
Q	Technician Code: This code is dedicated to a technician who meets the customer.
R	Number of Visits: Number of times that the technician visits the customer to solve the problem.
S	Row of Visit: The times that the technician visits the customer.

Column	Description
T	Date of Visit: The date at which the technician visits the branch.
U	Time of Visit: The time at which the technician visits the branch.
V	Technician Leaving Time: The time at which the technician finishes his/her task in the branch.
W	Customer Waiting Days: Number of days between the customer's call and technicians follow-up.
X	Piece Request Number: The warehouse is responsible to dedicate a number to the piece, which is requested by a technician.
Y	Code of Requested Piece: Code of the requested piece is written by a technician in this column.
Z	Serial Number of the Piece: Each piece has a specific serial number.
AA	Recording Date of the Information of the Requested Piece: The date that the technician requests the piece(s).
AB	Verification Date: Verification date is the date that the company verifies the information.
AC	Date of Piece Purchase: This column includes the date of a purchased piece to be substituted.
AD	Supply Location: The supply location is probably in capital of Iran (Tehran) or other provinces.
AE	Sent Date to a City: The sent date to cities is written in this column.
AF	Register Date in the Province Warehouse: The date of sending a new piece registered by a warehouse.
AG	Serial Number of Damaged Piece: It indicates the serial number of a damaged piece, which has been detached.
AH	Registering Date of Damaged Piece in System: The date of sending the detached pieces, which have been damaged, to the company.
AI	ATM Repair Technician: Is the person who fixes the detached piece.
AJ	Repair Description: Repairing tasks are written in this column.
AK	Date of Repair: The date of damaged piece repair.
AL	Code of the Used Piece for Repair: Sometimes a sub-piece- with a specific code- is used to repair the damaged piece.
AM	Date of Quality Rejection: After the repairing, the piece will be sent to Quality Control and may be rejected.
AN	Date of Quality Verification: If the repaired piece passes the QC test, the verification date will be registered in this column.
AO	Code of Quality Control Person: It is a code dedicated to a person who controls the quality.
AP	Voice of Customer Operator: It is a code dedicated to a person who makes calls with the customers.
AQ	Person Who Responds to the Voice of Customer: It includes name of the person who responds to the operator from the company.
AR	Causes of Dissatisfaction: After the operator makes a phone call, causes of customer dissatisfaction are written here.
AS	Special Piece: It is related to pieces sent to technician, but they did not work. Therefore, they return them to a warehouse as a special piece.

4.3. Data Preparation

It is quite common that users of the system enter data with errors or unusual values, even data might be stored in an inconsistent format with analysis objectives, and therefore, data must be prepared before mining. In other words, the data wished to be analyzed in the real world by data mining techniques are incomplete (lacking attribute values or certain attributes of interest, or containing only aggregate data), noisy (containing errors, or outlier values that deviate from the expected), and inconsistent (e.g., containing discrepancies in the department codes used to categorize items) (Han et al., 2011).

In this research, noises and inconsistency data was dealt with as following:

- ✓ Noisy data: omitted due to the small number of such records or lack of information to clean others.
- ✓ Inconsistent data: numerical and alphabetical pieces of information were coded in suitable numerical sets or ranges type.

Afterwards, the number of columns and rows were reduced to 3242 and 12 respectively. Table 3 indicates remained features and used conversions.

Table 3
Selected features and description of their preprocessing

No.	Feature	Preparation
1	Customer's Name	Rewritten by using English letters
2	Machine Code	Coded by using English letters
3	Province	Rewrite by using English letters
4	Grade	Coded by using English letters
5	Log Status	Coded by using English letters
6	Problem Code	Classification
7	Machine Status	Coding by using English letters
8	Technician Code	Classification
9	Number of Visits	Classification
10	Repair Time	Feature making and Classification
11	Code of Requested Piece	Classification
12	Supply Location	Coded by using English letters

4.4. Modeling

Based on the research framework, K-means algorithm was used to cluster data. In this paper, four indexes including DI (Dunn Index), (DBI) Davis-Bouldin Index, (SSE) Sum of Squared Error and Silhouette index have been used to determine the number of optimum clusters. Finally, number 2 was identified as the optimum number of clusters. Calculating each index for 2 to 10 clusters, the values were standardized. Tables 4-6 indicate the way through which, number of optimum clusters was calculated.

Table 4 shows the value of each index for 1 to 10 clusters.

Table 4.
Raw number of indexes for 1 to 10 clusters

Number of Clusters	Dunn	Davies-Bouldin	Sum of Squared Errors	Silhouette
1	0	0	1.64E+11	0
2	0.6964	0.5384	8.87E+09	0.7771
3	0.5399	0.6038	2.27E+09	0.7165
4	0.5888	0.565	3.82E+08	0.7997
5	0.3531	0.6218	1.92E+08	0.7666
6	0.2523	0.7076	1.23E+08	0.7302
7	0.4457	0.6431	6.38E+07	0.7526
8	0.3879	1.2093	3.06E+07	0.7039
9	0.2403	1.4325	2.32E+07	0.6646
10	0.219	1.3602	2.12E+07	0.6373

Calculating each index for 2 to 10 clusters, the values were standardized. Table 5 indicates standardized values of each index.

Table 5
Standardized values of each index

Number of clusters	Dunn	Davies-Bouldin	Sum of Squared Errors	Silhouette
2	1	1	1	0.8608
3	0.6722	0.9269	0.1614	0.488
4	0.7746	0.9703	0.2793	1
5	0.2809	0.9067	0.0512	0.7964
6	0.0697	0.8108	0.0272	0.572
7	0.4748	0.8829	0.0477	0.71
8	0.3537	0.2497	0.0571	0.4104
9	0.0445	0	0.0132	0.1682
10	0	0.0809	0	0
Number of selected clusters	2	2	2	4

As demonstrated in Table 5, the numbers of suggested clusters by each index is shown in the last row. These numbers of clusters are related to the maximum value of each index after standardization. Finally, the values of four indexes were added in each row (2 to 10 clusters), and represented in Table 6.

Table 6
Accumulated values of four indexes

Number of Cluster	Multi Criteria Decision Making	Ranking of Optimal Cluster Number
2	3.8608	1
3	2.2485	3
4	3.0243	2
5	2.0353	5
6	1.4797	6
7	2.1155	4
8	1.0709	7
9	0.2258	8
10	0.0809	9

Multi Criteria Decision Making column in Table 6 indicates the consensus of indexes on the number of clusters. That is to say, as this number becomes greater, consensus of indexes on the desired number of clusters becomes larger. As Ranking of Optimal Cluster Number column in Table 6 shows, indexes reach consensus first on two, then on four, and likewise. Consequently, 2 was selected as the optimum number of clusters for K-means algorithm. the feature values and their percent in each cluster are shown in Table 7.

Table 7
Feature values and their percent for each cluster

Features	Cluster 1		Cluster 2	
	Value	Percent	Value	Percent
Log Status	Normal	98.95	Normal	98.26
Supply Location	City Warehouse	70.62	City Warehouse	64.41
Code of Requested Piece	Electromechanical	66.32	Electromechanical	87.24
Customer's Name	Ansar Bank	76.84	Maskan Bank	98.87
Grade	Normal Transaction	58.13	High Transaction	51.65
Machine Code	3901427	97.03	3901427	83.77
Machine Status	No Guarantee, Contract with the Piece	99.62	No Guarantee, Non-Contractual	63.19
Number of Visits	Medium (2 times)	85.26	Medium (2 times)	44.62
Problem Code	Very Important	69.71	Very Important	84.64
Repair Time	Standard (greater than or equal to 2 hours and less than 4 hours)	52.92	Very Good (less than 2 hours)	83.07

Table 8 shows feature values and their percentage along with their importance in each cluster.

Table 8
Feature values and their percent for each cluster

	Cluster-1 	Cluster-2 	Importance >= 0.95 >= 0.90 < 0.90 Unknown
Code of Requested Piece	MECH 1.96%; ELME 66.32%; ELEC 31.72%	MECH 2.95%; ELME 87.24%; ELEC 9.81%	Important 1.00
Customer's Name	Refah 0.24%; Maskan 22.92%; Ansar 76.84%	Refah 1.13%; Maskan 98.87%; Ansar 0	Important 1.00
Grade	GA 37.85%; GB 58.13%; GD 4.02%	GA 51.65%; GB 43.32%; GD 5.03%	Important 1.00
Log Status	L0 1.05%; L1 98.95%	L0 1.74%; L1 98.26%	Marginal 0.90
Machine Code	MCA 1.96%; MCB 97.03 %; MCC 0; MCD 0; MCF 0.72%; MCG 0.29%	MCA 2.26%; MCB 83.77 %; MCC 0.09%; MCD 0.26%; MCF 6.77%; MCG 6.86%	Important 1.00
Machine Status	MA 0; MB 99.62%; MC 0.38%; ME 0	MA 63.19%; MB 28.99%; MC 7.73%; ME 0.09%	Important 1.00
Number of Visits	ID 6.46%; ME 85.26%; VW 3.21%; WE 5.07%	ID 8.51%; ME 44.62%; VW 16.23%; WE 30.64%	Important 1.00
Problem Code	A 69.7%; B 2.2%; C 0.05%; D 0.19%; E 27.8%	A 84.6%; B 12%; C 0.17%; D 0.78%; E 2.34%	Important 1.00
Repair Time	VG 43.06%; ST 52.92%; LO 2.58%; VL 1.44%	VG 83.07%; ST 10.76%; LO 3.47%; VL 2.69%	Important 1.00
Supply Location	SA 70.62%; SB 29.38%	SA 64.41%; SB 35.59%	Important 1.00

The next stage, Association Rules and GRI algorithm were used to have a more precise analysis of data. As it is shown in Table 9, the algorithm was run with different antecedents and consequents to get the meaningful rules. Lift index, Support, and Confidence were used for each rule to select strong rules. Table 9 shows some of the strongest rules, indicating a meaningful relationship between data.

Table 9
Extracted relationships using GRI algorithm

No.	Consequent	Antecedent	Support %	Confidence %	lift
1	Problem code: Important	Machine Code: 3901427 Province: Lorestan Grade: normal transaction Log Status: normal Number of Visits: weak (3 times) Code of Requested Piece: Electromechanical	1.4	100	16.885
2	Problem code: Very Important	Customer’s Name: Ansar Bank Machine Code: 3901427 Province: Golestan Grade: normal transaction Code of Requested Piece: Electromechanical	9.36	100	1.327
3	Problem code: Very Unimportant	Customer's Name: Ansar Bank Machine Code: 3901427 Province: Lorestan Grade: high transaction Log Status: normal Number of Visits: medium (2 times)	n	98.67	5.416
4	Repair Time: long (greater than or equal to 6 hours)	Customer Name: Maskan bank Machine Code: 3901427 Grade: low transaction Problem Code: very important Machine Status: No guarantee, Non-contractual Technician Code: TZT	0.45	100	46.083
5	Repair Time: very good (less than 2 hours)	Customer’s Name: Ansar Bank Machine Code: 3901427 Grade: normal transaction Problem Code: very important Machine Status: No guarantee, contract with piece Technician Code: TZH Number of Visits: medium (2 times)	9.31	100	1.75
6	Repair Time: long (greater than or equal to 4 hours and less than 6 hours)	Customer’s Name: Maskan Bank Machine Code: 3901427 Problem Code: very important Machine Status: No guarantee, contract with piece Technician Code: TZR Code of Requested Piece: Electromechanical	0.59	100	35.677

No.	Consequent	Antecedent	Support %	Confidence %	lift
7	Repair Time: very long (greater than or equal to 6 hours)	Machine Code: 3901427 Grade: high transaction Problem Code: very important Technician Code: TYB Number of Visits: ideal (1 time) Code of Requested Piece: Electromechanical	0.41	100	46.083
8	Number of Visits: weak (3 times)	Machine Code: 3901427 Grade: normal transaction Problem Code: very important Technician Code: TZX	7.05	100	6.87
9	Number of Visits: weak (3 times)	Machine Code: 3901427 Grade: normal transaction Problem Code: important Technician Code: TZA Repair Time: very good (less than 2 hours) Supply Location: City Warehouse	1.36	100	6.87
10	Number of Visits: Very Weak (4 to 7 times)	Machine Code: 3901427 Grade: high transaction Problem Code: very important Technician Code: TZQ Repair Time: very good (less than 2 hours) Supply Location: City Warehouse	5.47	96.69	12.363
11	Code of Requested Piece: Electromechanical	Customer's Name: Ansar Bank Machine Code: 3901427 Province: Bushehr Problem Code: very important	1.08	100	1.356

Finally, to predict and suggest solutions to improve services, the classification method has been used. C5.0, C&R3, CHAID, Bayes, Neural Network, Support Vector Machine and Logistic are the classification algorithms used in this research. Using these algorithms, different targets and inputs were investigated and their accuracy was calculated. The results indicated that Support Vector Machine and C5.0 had the highest rates of accuracy respectively. Considering the fact that extracting rules were not defined in Support Vector Machine, C5.0 algorithm, which is based on Decision Tree, was selected to classify the data. In fact, how the rules are extracted from dataset is clear in C5.0 algorithm .

Table 10 draws a comparison among the accuracy of algorithms.

Table 10
Comparison among the accuracy of algorithms

Target	Inputs	Algorithm	Accuracy %	Ranking of Accuracy
Problem Code	Customer’s Name Machine Code Province Grade Code of Requested Piece Number of Visits Log Status	C5.0	91.84	2
		C&R Tree	89.81	3
		CHAID	89.81	3
		Bayes	89.13	4
		Neural Network	83.01	6
		Logistic	85.15	5
		SVM	93.15	1

After the algorithm started operating in each mode, the importance of the features became clear. Features with the lowest importance were deleted one by one, and each time the algorithm ran again to investigate if the results are the same, omit the inputs with low importance. The extracted rules are as follows:

1. If Number of Visits = ME (medium (2 times)), Province = East Azarbayjan and Customer's Name = Maskan Bank, then Problem Code = B (very important). 1.9 percent of data belongs to this group.

To elucidate it, Fig 2 indicates the stages of extracting this rule.

2. If Customer's Name = Maskan Bank, Number of Visits = VW (very weak (4 to 7 times)), Technician Code = TZQ (Fars) and Problem Code = A (very important), then Repair Time = VG (very good (less than 2 hours)). 5.28 percent of data belongs to this group.

3. If Grade = low transaction, Technician Code = TZT (North Khorasan) and Problem Code = A (very important), then Repair Time = VL (very long (greater than or equal to 6 hours)). 0.45 percent of data belongs to this group.

4. If Machine Code = MCB (3901427), Repair Time = VG (very good (less than 2 hours)), Problem Code = A (very important) and Technician Code = TZR (Khorasan E Razavi), then Number of Visits = VW (very weak (4 to 7 times)). 1.49 percent of data belongs to this group.

5. If problem code = A (very important), Repair Time = VG (very good (less than 2 hours)), Grade = GA (high transaction) and Technician Code = TZQ (Fars), then Number of Visits = VW (very weak (4 to 7 times)). 5.51 percent of data belongs to this group.

6. If Machine Status = MB (No Guarantee, Contract with the Piece), Customer's Name = Maskan Bank and Province = Markazi, then Code of Requested Piece = ELEC (electric). 6.64 percent of data belongs to this group.

4.5. Model Evaluation and Deployment

Stages 5 and 6 are explained together in this section. As it is indicated in Table 7, data was divided into two groups by K-means algorithm. Each cluster is explained as follows:

Cluster number 1: this cluster includes 2090 data. Almost all data in this cluster indicates Machine Status without guarantee – Pieces Contract, Normal Status, and Machine Code 3901427. The most share among all requested pieces belongs to pieces, which their Code of Requested Piece is Electromechanical, Customer's Name is Ansar Bank, Supply Location is City Warehouse, Problem Code is very important, Number of Visits is Medium (2 times), and Grade is Normal. In fact, the allocated values for each feature make more than half of the data in this cluster. Moreover, almost half of data of this group includes Standard Repair Time (more than or equal to 2 hours and less than 4 hours).

Cluster number 2: this cluster includes 1152 data. Almost all of the data indicates normal Log Status and the Customer’s Name is Maskan Bank. For more than half of the data, Supply Location is City Warehouse, Code of Requested Piece is Electromechanical, Code of ATM is 3901427, Machine Status is without guarantee, Non-contractual, Problem Code is very important, and Repair Time is Very Good (less than 2 hours). In this cluster, almost half of the transaction values show hightransaction and Number of Visits is Medium (2 times) includes less than half of the data.

In the next stage, GRI algorithm was deployed to conduct a more accurate analysis, while discovering the rules among the data, rules were evaluated by Lift index, Support and Confidence. According to Table 9, it was concluded that reducing the antecedents of rules in rows 2, 3, 4, and 5 would lead to results that are more desirable. Accordingly, Table 11 shows the results after antecedent reduction.

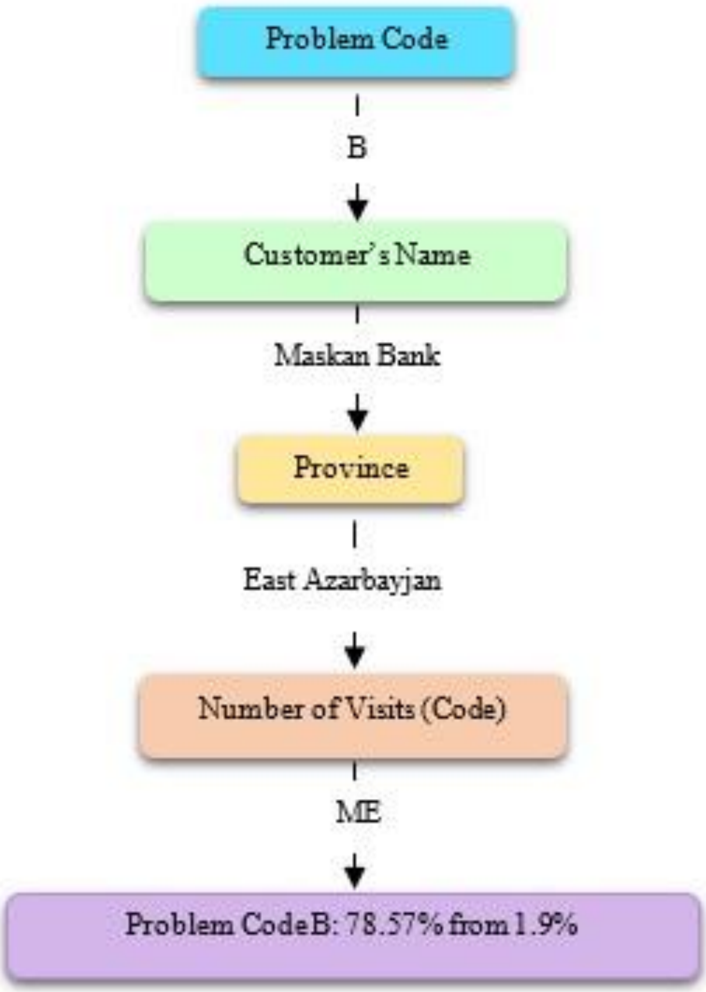

Fig 2. The procedure of extraction rule number one

Table 11
Extracted relations using GRI algorithm after antecedent reduction

No.	Consequence	Antecedents	Support %	Confidence %	Lift
1	Problem Code: very important	Customer's Name: Ansar Bank Machine Code: 3901427 Province: Golestan	15.18	91.46	4.869
2	Problem Code: very unimportant	Customer's Name: Ansar Bank Machine Code: 3901427 Province: Lorestan	8.76	99.3	1.324
3	Repair Time: very long (greater than or equal to 6 hours)	Machine Code: 3901427 Problem Code: very important Technician Code: TZT	0.46	86.67	46.063
4	Repair Time: very good (less than 2 hours)	Machine Code: 3901427 Problem Code: very important Technician Code: TZH	8.73	99.29	1.733

As it is clarified in rows 1 and 2 in Table 11, different provinces are struggling with different ATM repair problems, that could be due to different weather conditions, culture of using ATMs, and likewise. Rows 3 and 4 indicate that technicians perform differently regarding Repair Time that could be due to differences in technicians' experiences, training, and such. Through identifying and resolving effective factors, the severity of problems could be decreased, while repair time could be minimized as well, and the customers' satisfaction would increase consequently.

Finally, C5.0 method has been used for data classification as a predictive analytics in data mining. Extracted rules from classification can be used in prediction. Carrying out a review on the mentioned rules, the situations leading to important problems, weak Number of visits, and such were realized. Moreover, two rules was found, which are common in GRI and C5.0 algorithms by investigating the extracted rules. Table 12 indicates these rules.

Table 12
GRI and C5.0 algorithms common rules

	Association rules					Classification	
No.	**Lift**	**Confidence %**	**Support %**	**Antecedent**	**Consequent (target)**	**Inputs**	**Number of data %**
1	46.08	100	0.45	Customer’s Name: Maskan Bank Machine Code: 3901427 Grade: low transaction Problem Code: very important Machine Status: No guarantee, Non-contractual Technician Code: TZT	Repair time: very long (greater than or equal to 6 hours)	Grade: low transaction Problem Code: very important Technician Code: TZT	0.452
2	12.36	96.69	5.47	Machine Code: 3901427 Grade: high transaction Problem Code: very important Technician Code: TZQ Repair Time: very good (less than 2 hours) Supply Location: City Warehouse	Number of referrals: very weak (4 to 7 times)	Grade: high transaction Problem Code: very important Technician Code: TZQ Repair Time: very good (less than 2 hours)	5.515

After evaluating the model and defining useful rules, a report was conducted. The report provided practical information such as rules shown in Table 12. Therefore, it helped company's managers to use its information in their decision-making process and formulating strategies.

5. Results

According to Tables 10-12, there can be seen some extracted rules from data sets, using GRI and C5.0 algorithms. Tables 11 and 12 are indicative of the rules, which have been either extracted through reduction of antecedents, or they are common in two algorithms. Therefore, rules of these two tables are more reliable. Rows 1 and 2 of Table 11 clearly state that the same Machines’ Code and the same customers in different provinces caused significantly different Problems' Code that can be due to differences in weather condition, culture of using ATMs, and likewise. Thus, it can be said one can investigate the way an ATM is being used and prevent relevant problems. Considering the weather condition, the company can prevent selling ATM 3901427 in Golestan Province and substitute similar one suitable for that province. Additionally, rows 3 and 4 indicate the same Machine's Code and the same Problems' Code; notably there can be seen that different technicians seemed to record significantly different Repair Times that could be due to their training background level of their expertise and such. To resolve this problem, the technicians' level of expertise in South Khorasan Province should be investigated so that the factors causing this matter could be identified. Row 2 of Table 12 identifies the major factors leading to weak Number of Visits (4 to 7 times). ATM 3901427, High Transactions, Very Important Problem, the technicians of Fars Province, Supplying ATM pieces by the City Warehouse, and a Good Repair Time (less than 2 hours) led to the very weak number of visits. With respect to Very Important Problem, High Transaction, and Supplying ATM pieces by the City Warehouse, the Very Weak Number of Visits is reasonable. On the contrary, no matter what causes this result, the Number of Visits should be minimized. In other words, although the intensity of a problem, remote warehouses, and the level of technicians' competency cannot justify the Number of Visits, effort should be made to consider these culprits and minimize the Number of Visits.

6. Conclusion

In this article, CRISP-DM method was used to carry out the research stages by using the data of an ATM manufacturing company. After data understanding, it was realized that there were many missing values and unusable features. Therefore, in preparation stage, a large amount of data was omitted, while the remained one was coded and classified. After preparing data for entering in Clementine, first clustering and K-means algorithm were used as a descriptive method. Determining the optimum cluster has been an important step in K-means algorithm, which is conducted by indexes such as Dunn, Silhouette, Davies-Bouldin, and Sum of Squared Errors (SSE). Accordingly, the optimum number of clusters was defined 2. After using K-means and dividing the data into two groups, GRI algorithm was deployed to conduct an accurate analysis, while discovering the relations among data. To evaluate extracted rules, Lift index, Support, and Confidence were used. Finally, classification was done as a predictive method. To use an appropriate algorithm in this article, the accuracy of algorithms such as C5.0, C&R Tree, CHAID, Bayes, Neural Network, Logistic, and Support Vector Machine were measured. The results clearly indicated that Support Vector Machine and C5.0 had the highest rates of accuracy respectively. Considering the fact that C5.0 algorithm is strongly based on Decision Tree, C5.0 was selected. The extracted rules by this method are greatly helpful for predictions of organizations. Additionally, there were two extracted rules in C5.0 and GRI, which were in common. Regarding future researches, other algorithms of Clustering, Association Rules, and Classification can be used to compare the obtained results with the current research's result. Moreover, Fuzzy Clustering can be deployed to cluster different pieces.

6.1. Ideas for Further Research

(i) Other clustering, association rules, and classification algorithms could be considered, and their results could be compared with algorithms used in this work.

(ii) Regarding the results of clustering, it is suggested to use fuzzy clustering to obtain results that are more accurate.

(iii) Technicians' Work experience, age, and training hours may have profound effects on their performance. Unfortunately, dataset, which was studied in this work, did not contain this information. It is firmly recommended that this information should be considered in further research.

Acknowledgments

The authors extend their appreciation to Ms. Masoudi, from the ATM Company, for her valuable guidance and supports during this work entire process.

References

Ahmed, S.R. (2001). Applications of data mining in retail business. *International Conference on Information Technology: Coding and Computing, 2004. Proceedings. ITCC* 2004, IEEE, 455-459.

Arumawadu, H.I., Rathnayaka, R.K.T., & Illangarathne, S. (2015). Mining Profitability of Telecommunication Customers Using K-Means Clustering. *Journal of Data Analysis and Information Processing*, 3(3), 63.

Balaji, S., & Srivatsa, S. (2012). Customer segmentation for decision support using clustering and association rule based approaches. *International Journal of Computer Science & Engineering Technology*, 3(11), 525-529.

Berry, M.J., & Linoff, G. (1997). Data mining techniques: for marketing, sales, and customer support. *John Wiley & Sons*, Inc.

Bhuvaneswari, C., Aruna, P., & Loganathan, D. (2014). A new fusion model for classification of the lung diseases using genetic algorithm. *Egyptian Informatics Journal*, 15(2), 69-77.

Chen, L-F. (2015). Exploring asymmetric effects of attribute performance on customer satisfaction using association rule method. *International Journal of Hospitality Management*, 47, 54-64.

Chen, Y-S., Cheng, C-H., Lai, C-J., Hsu, C-Y., & Syu, H-J. (2012). Identifying patients in target customer segments using a two-stage clustering-classification approach: A hospital-based assessment. *Computers in Biology and Medicine*, 42(2), 213-221.

Cheng, C-H., & Chen, Y-S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications*, 36(3), 4176-2184.

Chiang, W-Y. (2011). To mine association rules of customer values via a data mining procedure with improved model: An empirical case study. *Expert Systems with Applications*, 38(3), 1716-1122.

Cil, I. (2012). Consumption universes based supermarket layout through association rule mining and multidimensional scaling. *Expert Systems with Applications*, 39(10), 8611-8625.

Cimpoiu, C., Cristea, V-M., Hosu, A., Sandru, M., & Seserman, L. (2011). Antioxidant activity prediction and classification of some teas using artificial neural networks. *Food chemistry*, 127(3), 1323-1328.

Cremaschi, S., Shin, J., & Subramani, H.J. (2015). Data clustering for model-prediction discrepancy reduction–A case study of solids transport in oil/gas pipelines. *Computers & Chemical Engineering*, 81, 355-363.

da Silva, C.E.T., Filardi, V.L., Pepe, I.M., Chaves, M.A., & Santos, C.M.S. (2015). Classification of food vegetable oils by fluorimetry and artificial neural networks. *Food control*, 47, 86-91.

de Amorim, R.C., & Hennig, C. (2015). Recovering the number of clusters in data sets with noise features using feature rescaling factors. *Information Sciences*, 324, 126-145.

de Oña, J., de Oña, R., & Calvo, F.J. (2012). A classification tree approach to identify key factors of transit service quality. *Expert Systems with Applications*, 39(12), 11164-11171.

Dhandayudam, P., & Krishnamurthi, D.I. (2012). An improved clustering algorithm for customer segmentation. *International Journal of Engineering Science and Technology*, 4(2), 99-102.

Doub, A.E., Small, M.L., Levin, A., LeVangie, K., & Brick, T.R. (2016). Identifying users of traditional and Internet-based resources for meal ideas: An association rule learning approach. *Appetite*, 103, 128-136.

Elyasigomari, V., Mirjafari, M., Screen, H.R., & Shaheed, M.H. (2015). Cancer classification using a novel gene selection approach by means of shuffling based on data clustering with optimization. *Applied Soft Computing*, 35, 43-51.

Friedman, J., Hastie, T., & Tibshirani, R. The elements of statistical learning. (2001). *Springer series in statistics*, New York.

Giraud-Carrier, C., & Povel, O. (2003). Characterising data mining software. *Intelligent Data Analysis*,7(3), 181-192.

Golmah, V., & Mirhashemi, G. (2012). Implementing a data mining solution to customer segmentation for decayable products-a case study for a textile firm. *International Journal of Database Theory and Application*, 5(3), 73-90.

Günter, S., & Bunke, H. (2003). Validation indices for graph clustering. *Pattern Recognition Letters*. 24(8), 1107-1113.

Gurrutxaga, I., Muguerza, J., Arbelaitz, O., Pérez, J.M., & Martín, J.I. (2011). Towards a standard methodology to evaluate internal cluster validity indices. *Pattern*

Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. 3rd edition, *Elsevier*.

Hosseini, S.M.S., Maleki, A., & Gholamian, M.R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, 37(7), 5259-5264.

İşeri, A., & Karlık, B. (2009). An artificial neural networks approach on automobile pricing. *Expert Systems with Applications*, 36(2), 2155-2160.

Ivančević, V., Tušek, I., Tušek, J., Knežević, M., Elheshk, S., & Luković, I. (2015). Using association rule mining to identify risk factors for early childhood caries. *Computer methods and programs in biomedicine*, 122(2), 175-181.

Karabatak, M., & Ince, M.C. (2009). A new feature selection method based on association rules for diagnosis of erythemato-squamous diseases. *Expert Systems with Applications*, 36(10), 12500-12505.

Kareem, B., & Lawal, A. (2015). Spare parts failure prediction of an automobile under criticality condition. *Engineering Failure Analysis*, 56, 69-79.

Kargari, M., & Sepehri, M.M. (2012). Stores clustering using a data mining approach for distributing automotive spare-parts to reduce transportation costs. *Expert Systems with Applications*, 39(5), 4740-4748.

Keshavarzi, A., Sarmadian, F., Omran, E-SE., & Iqbal, M. (2015). A neural network model for estimating soil phosphorus using terrain analysis. *The Egyptian Journal of Remote Sensing and Space Science*, 18(2), 127-135.

Kim, K-j., & Ahn, H. (2008). A recommender system using GA K-means clustering in an online shopping market. *Expert systems with applications*, 34(2), 1200-1209.

Kuo, R.J., & Zulvia, F.E. (2018). Automatic clustering using an improved artificial bee colony optimization for customer segmentation. *Knowledge and Information Systems*, 1-27.

Lee, P-J., Hu, Y-H., & Lu, K-T. (2018). Assessing the Helpfulness of Online Hotel Reviews: A Classification-based Approach. *Telematics and Informatics*, 35(2), 436-445.

Luo, Y., Li, Z., Guo, H., Cao, H., Song, C., Guo, X., & Zhang Y. (2017). Predicting congenital heart defects: A comparison of three data mining methods. *PloS one*, 12(5), e0177811.

Mahdavi, I., Cho, N., Shirazi, B., & Sahebjamnia, N. (2008). Designing evolving user profile in e-CRM with dynamic clustering of Web documents. *Data & Knowledge Engineering*, 65(2), 355-372.

Manolopoulou, .E, Kotsiantis, S., & Tzelepis, D. (2015). Application of association and decision rules on intellectual capital. *Knowledge Management Research & Practice*, 13(2), 225-234.

Martin, I., & Bestle, D. (2018). Automated eigenmode classification for airfoils in the presence of fixation uncertainties. *Engineering Applications of Artificial Intelligence*, 67, 187-196.

Mirabadi, A., & Sharifian, S. (2010). Application of association rules in Iranian Railways (RAI) accident data analysis. *Safety Science*, 48(10), 1427-1435.

Mitra, S., Pal, S.K., & Mitra, P. (2002). Data mining in soft computing framework: a survey. IEEE transactions on neural networks, IEEE, 13(1), 3-14.

Molina, C., Stepien, O., Pessegue, B., & Rameau, J-P. (2016). PKOM: A tool for clustering, analysis and comparison of big chemical collections. *Digital Signal Processing*, 48, 1-11.

Movagharnejad, K., Mehdizadeh, B., Banihashemi, M., & Kordkheili, M.S. (2011). Forecasting the differences between various commercial oil prices in the Persian Gulf region by neural network. *Energy*, 36(7), 3979-3984.

Nasibov, E., Savaş, S.K., Vahaplar, A., & Kınay, A.Ö. (2016). A survey on geographic classification of virgin olive oil with using T-operators in Fuzzy Decision Tree Approach. *Chemometrics and Intelligent Laboratory Systems*, 155, 86-96.

Ngai, E.W., Xiu, L., & Chau, D.C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert systems with applications*, 36(2), 2592-2602.

Palomares-Salas, J.C., Agüera-Pérez, A., de la Rosa, J.J.G., & Moreno-Muñoz, A. (2014). A novel neural network method for wind speed forecasting using exogenous measurements from agriculture stations. *Measurement*, 55, 295-304.

Parvaneh, A., Abbasimehr, H., & Tarokh, M.J. (2012). Integrating AHP and data mining for effective retailer segmentation based on retailer lifetime value. *Journal of Optimization in Industrial Engineering*, 5(11), 25-31.

Rumpf, T., Mahlein, A-K., Steiner, U., Oerke, E-C., Dehne, H-W., & Plümer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture*, 74(1), 91-99.

Shaw, M.J., Subramaniam, C., Tan, G.W., & Welge, M.E. (2001). Knowledge management and data mining for marketing. *Decision support systems*, 31(1), 127-137.

Seret, A., vanden Broucke, S.K., Baesens, B., & Vanthienen, J. (2014). A dynamic understanding of customer behavior processes based on clustering and sequence mining. *Expert Systems with Applications*, 41(10), 4648-4657.

Shih, Y-Y., & Liu, D-R. (2008). Product recommendation approaches: Collaborative filtering via customer lifetime value and customer demands. *Expert Systems with Applications*, 35(1), 350-360.

Shim, B., Choi, K., & Suh, Y. (2012). CRM strategies for a small-sized online shopping mall based on association rules and sequential patterns. *Expert Systems with Applications*, 39(9), 7736-7742.

Silvera SAN, Mayne ST, Gammon MD, Vaughan TL, Chow W-H, Dubin JA, Dubrow, R., Stanford, J.L., West, A.B., Rotterdam, H. & Blot, W.J. (2014). Diet and lifestyle factors and risk of subtypes of esophageal and gastric cancers: classification tree analysis. *Annals of epidemiology*, 24(1), 50-57.

Sudha, L., Dillibabu, R., Srinivas, S.S., & Annamalai, A. (2016). Optimization of process parameters in feed manufacturing using artificial neural network. *Computers and Electronics in Agriculture*, 120, 1-6.

Tormo, M.T., Sanmartin, J., & Pace, J. F. (2009). Update and improvement of the traffic accident data collection procedures in Spain: The METRAS method of sequencing accident events. *In 4th IRTAD Conference*. Seoul, Korea. 16-17.

Tsai, C-F., & Chen, M-Y. (2010). Variable selection by association rules for customer churn prediction of multimedia on demand. *Expert Systems with Applications*, 37(3), 2006-2015.

Tsiptsis, K.K., & Chorianopoulos, A. (2011). Data mining techniques in CRM: inside customer segmentation, *John Wiley & Sons*.

Vafaie, M., Ataei, M., & Koofigar, H.R. (2014). Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. *Biomedical Signal Processing and Control*, 14, 291-296.

Van Der Schaaf, T., & Kanse, L. (2004). Biases in incident reporting databases: an empirical study in the chemical process industry. *Safety Science*. *42*(1), 57-67.

Venkatesh, K., Ravi, V., Prinzie, A., & Van den Poel, D. (2014). Cash demand forecasting in ATMs by clustering and neural networks. *European Journal of Operational Research*, 232(2), 383-392.

Vermeulen-Smit, E., Ten Have, M., Van Laar, M., & De Graaf, R. (2015). Clustering of health risk behaviours and the relationship with mental disorders. *Journal of affective disorders*, 171, 111-119.

Videla-Cavieres, I.F., & Ríos, S.A. (2014). Extending market basket analysis with graph mining techniques: A real case. *Expert Systems with Applications*, 41(4), 1928-1936.

Vijayarani, S., & Sudha, S. (2015). An efficient clustering algorithm for predicting diseases from hemogram blood test samples. *Indian Journal of Science and Technology*, 8(17).

Wang, J-Y., Liu, C-S., Lung, C-H., Yang, Y-T., & Lin, M-H. (2017). Investigating spousal concordance of diabetes through statistical analysis and data mining. *PloS one*, 12(8), e0183413.

Wang, B., Miao, Y., Zhao, H., Jin, J., & Chen, Y. (2016). A biclustering-based method for market segmentation using customer pain points. *Engineering Applications of Artificial Intelligence*, 47, 101-109.

Weng, J., Zhu, J-Z., Yan, X., & Liu, Z. (2016). Investigation of work zone crash casualty patterns using association rules. *Accident Analysis & Prevention*, 92, 43-52.

Zhang, J., Feng, Q., Zhang, X., Zhang, X., Yuan, N., Wen, S., Wang, S., & Zhang, A. (2015). The use of an artificial neural network to estimate natural gas/water interfacial tension. *Fuel*, 157, 28-36.

Zhao, Z., Yang, Z., Lin, H., Wang, J., & Gao, S. (2016). A protein-protein interaction extraction approach based on deep neural network. *International Journal of Data Mining and Bioinformatics*, 15(2), 145-164.

This article can be cited: Bagherighadikolaei, S. P. Ghousi, R. & Haeri,A. (2021). A Data Mining Approach for Forecasting Failure Root Causes: A Case Study in an Automated Teller Machine (ATM) Manufacturing Company. *Journal of Optimization in Industrial Engineering*. 13 (2), 101-121.

http://www.qjie.ir/article_671443.html

DOI: 10.22094/JOIE.2020.1863364.1630