PREDICTION OF CREDIT CARD PAYMENT NEXT MONTH THROUGH TREE NET DATA MINING TECHNIQUES

Ahmed Mohammed Hussein [1)], Hadeel Qasem Gheni [1)], Wed Kadhim Oleiwi [1)], Zahraa Yaseen Hasan [2)]

[1)] Department of Computer Science, College of Science for Women (SCIW), University of Babylon, Babylon, Iraq.
wsci.ahmed.mohammed@uobabylon.edu.iq, wsci.hadeel.qasem@uobabylon.edu.iq,
wsci.wed.kadhum@uobabylon.edu.iq
[2)] Department of Laser Physics Science, College of Science for Women (SCIW), University of Babylon, Babylon, Iraq.
Zahraaayaseen395@yahoo.com

Abstract: A number of research initiatives have recently been launched around the world regarding the conceptualization, specification, design and development principles of the future use of credit cards, storing secret information on them, while most time we use them for online payment. In addition, if it has enough money, we can pay for what we need at any time. Therefore, the goal of this proposed research is to use data mining techniques to predict credit card payment next month. Our proposed system contains five steps: (a) find the suitable database from the internet because this database is not available in Iraq, (b) pre-process the credit card database based on person correlation matrix to determine which feature is less correlated with other to remove it and reduce the time of prediction, (c) split pre-processing database into two parts training and testing dataset, (d) apply TreeNet prediction data mining techniques (TPDMT) on training dataset to test if we need payment next month or do not, find the optimal tree. TreeNet based on Boosting Machine usually makes the predictor to use Decision Trees (DTs). (e) Finally, pass the testing dataset on the optimal tree results from TPDMT, then using the five measures related to confusion matrix to evaluate the results including "Accuracy (AC), recall or true positive rate (TP), precision (P), F-measure (considers both precision and recall) and Fb".

1. INTRODUCTION

Without cash, credit card is better and more suitable way to sell services. Credit card is a small card used for payments and purchases; many opportunities for electronic commerce are available because of the ease of payment over the Internet [4]. In general, there are two types of it: physical and virtual [3]. Physical card such as (American Express, Discover, MasterCard, Visa, etc…), which owner must show during the purchase process. While virtual card is online, it is used in free services offered by the original card issuer. Credit cards consider critical to improving the economic power of any country [1]. Using it by an ordinary citizen is a major step to promoting a highly desirable non-monetary economy [2]. In recent decades and with

the rapid development of the globalization of the economy, credit cards have become the most popular in business dealings. [5]. In everyday life, buying goods, services and trading online has increased [3]. These increases made payments more efficient, cancel cash handling fees, reduce the risk of theft and increase economic activities [1]. The use of a credit card often leads to adjustment of relationships between the power of money and distrust and anxiety [20]. Finally, the use of the credit card is necessary where the holder does not need to carry large amount of money with him and it provides the possibility of payment using any kind of currency, thus, does not require currency exchange by the customer.

Data mining is calculations to analyse large-scale data so as to extract useful patterns and

information [6]. A wide range of data is transforming into knowledge via data mining, which can be seen as a result of the natural development of information technology [7]. Data is a tool like any other tool and it is not important to know how it works but it is necessary to know how it will be used [18]. It works with two policies, supervised (with training set) and unsupervised (without training set) learning [10], the modelling objectives most commonly used in data mining are prediction and classification [11]. The success of data mining in many fields including electronic business, shopping and sales led to its use in other applications [9]. Nowadays, data mining has a major role in many applications like business administration, government sides, engineering and scientific fields, healthcare and educational foundations [8].

In data mining, the most common methods is finding rules for prediction from empirical data groups [14]. Prediction is the process of results prediction based on existing data, relying on the relationship between something you know and something we need to predict. In general, there are three kinds of prediction: classification, density estimation and regression [12]. Classification and regression can be distinguished according to the type of prediction, where the output of the classification is a categorical class while the model in the regression learns continuous values [13]. This type of data mining can help business leaders make better decisions, which may be beneficial by knowing many of the statistics required in the future. Prediction used in data mining to rationalize decisions in many areas, such as banks, build future predictions and explore behaviour and trends, allowing accurate judgment to be made and taken in a timely manner.

The remainder of the research is organizing as follows. The related works are presented in Section 2. The techniques used in prediction are presented in Section 3. The suggested methods are used in designing the new predicator in Section 4. Generating steps of the proposed predicator and challenges are shown in Section 5. The experiments are described in Section 6. In the end, Section 7 includes discussion and conclusion of the research.

2. LITERATURE REVIEW

A lot of previous works have used data mining methods to deal with credit card applications and problems. One of the most popular problem is fraud detection; it has taken a lot of effort due to the increasing number of frauds all over the world.

In 1999, Philip Chan and other researchers proposed a method of combining multiple learned fraud detectors under a "cost model". The outcomes show that we can fundamentally diminish misfortune because of misrepresentation through circulated information mining of extortion models [19].

In another time, Yap Bee Wah and Irma Ibrahim showed in their paper the development and correlation of three credit scoring models: logistic relapse (LR) model, arrangement and relapse tree (CART) model and neural system (NN) model to segregate amongst rejected and acknowledged charge card candidates of a bank [20].

In 2011, V. Bhusari and S. Patil showed how Hidden Markov Model (HMM) used to discover the fraud transaction of credit card with least false alarm. A HMM based framework is at first concentrated on spending profile of the cardholder and took after by checking an approaching exchange against spending conduct of the cardholder, on the off chance that it is not acknowledged by their proposed HMM with adequate likelihood, then it would be a fake exchange.

In the same year, Raghavendra Patidar and Lokesh Sharma tried to detect fraudulent transaction through the neural network along with the genetic algorithm. Genetic algorithm is used for setting the system topology option, shrouded layers numbers, hubs numbers as a part of the neural system configuration for the card payment misrepresentation identification. For the learning reason, to simulate neural system they utilized directed learning nourish forward back proliferation calculation. At long last they propose some future work that should be possible in making extortion recognition.

In 2013, Hetvi Modi proposed a system that is capable of detecting fraud by taking into account the spending routine of credit card holders without its importance. Ordinarily, when the points of interest of things obtained in individual exchanges are unknown this system is not a system of Fraud Detection. The proposed framework will be a perfect decision for tending to this issue of current extortion location framework. Another vital preferred standpoint of proposed framework will be an extraordinary decrease in the quantity of False Positives exchanges. FDS module of proposed framework will get the card points of interest and the estimation of procurement to confirm, whether the exchange is honest to goodness or not. In the event that the Fraud Detection System module will affirm the exchange to be of misrepresentation, it will raise an alert, and the exchange will be declined.

Marcos and Sousa and Reginaldo Figueiredo, 2014, developed models to analyse the capacity of a credit union's members to settle their commitments, using a decision tree—C4.5 algorithm and an artificial neural network—multilayer perceptron

algorithm. It is concluded that for the problem of credit analysis, the models have statistically similar results and may aid in a cooperative's decision-making process.

Another research in 2014 used Hidden Markov Model (HMM), model the order of service while making a credit card transaction. If the qualified HMM does not accept an incoming credit card transaction with sufficiently high probability, this will be deemed fraudulent.

Jae Kwon Bae1 and Jinhwa Kim, 2015 build many of personal credit rating prediction models based on the UDM and benchmarks their performance against other models which employ logistic regression (LR), Bayesian style frequency matrix (BFM), multilayer perceptron (MLP), classification tree methods (C5.0), and neural network rule extraction (NR) algorithms. To verify the feasibility and effectiveness of UDM, personal credit data and personal loan data provided by a Financial Holding Company (FHC) were used in this study. Empirical results indicated that UDM outperforms other models such as LR, BFM, MLP, C5.0, and NR.

Mrs. Poonam and other researchers, 2016, give suggestions for a new fraud detection technique in credit cards which can be implemented and which will seize the essence of the existing techniques and may combine few of them to give superior fraud detection tool.

3. PREDICTION DATA MINING TECHNIQUE

In this part, the main principle used to identify and resolve the problem will be discussed.

3.1 TREENET ALGORITHM

One of the most revolutionary developments in data mining technologies is TreeNet, undertaken by one of the most famous researchers in data mining, Jerome Friedman. TreeNet is designed for predictive modelling with very high accuracy. TreeNet tries to achieve this aim even if the required models are too complex, so models are perhaps relatively hard to understand in detail. In general, TreeNet model is made up of small trees ranging from several dozen to several hundred, usually each of them does not exceed the size of two to eight terminal points. Spiritually, this model is identical to the long series expansion (such as a Fourier or Taylor's series) - A set of factors that gradually become more accurate with continued expansion [12]. The model can be considered as a serial expansion that approximates the real functional relationship.

$$F(x) = F0\beta1T1(X) + \beta2T2(X) + \cdots + \beta MTM(X), \quad (1)$$

where each TM is a small tree.

This section presents techniques for improving the classification accuracy by aggregating multiple classifier predictions. Such approaches are methods known as the ensemble or classifier combination. An ensemble approach generates a collection of base classifiers from training data and performs classification by taking a vote on each base classifier's predictions.

Generally, two substantial prerequisites must meet for an ensemble classifier to be with a performance higher than the performance of the individual classifier: (1) the independence of the base classifiers from each other, (2) the performance of the base classifiers must be better than the performance of the classifier that performs random estimation.

The figure below illustrates the ensemble method as a logical view. Idea is centred on taking the original data and creating several classifiers from it and with the classification of unknown cases, their predictions are translated. The ensemble classifiers can be created in several strategies, [22].
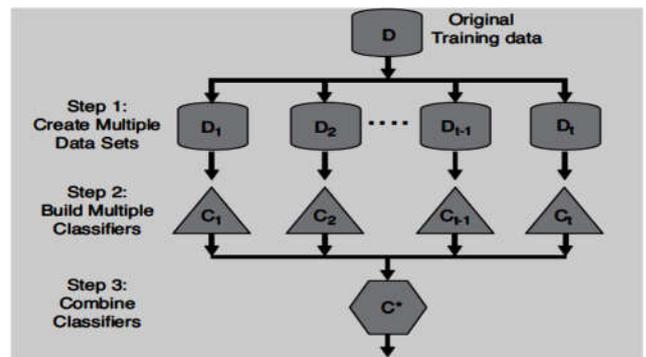

Figure 1 – A logical view of the ensemble learning method [24], [22]

When manipulating the training set, numerous training groups are formed in this way by reconstituting the original data according to some sample distributions. From each training group, a classifier is constructed by using a specific learning algorithm. Two examples of ensemble methods are Bagging and boosting which are working on manipulating their training sets. This work is based on using boosting learning algorithm in prediction. We can explain the main classification methods as follows:

- At the pre-processing stage select the important features from the total input features to each training set. These subsets of features are chosen randomly or based on the recommendation of domain experts.

- At the classification stage determine the label for each class. The classification method works with large number of classes.
- By this procedure, we can apply each of the learning algorithms several times on the same training data and this may result in different models. For example, TreeNet can generate different models by changing its number of trees or number of levels in each model.

The first three approaches are generic methods that are applicable to any classifiers, on another hand; the fourth approach depends on the type of classifier used. The base classifiers for most of these approaches can be generated sequentially (one after one) or in parallel (all at once). Algorithm bellow shows the steps needed to build an ensemble classifier in a sequential manner [23], [22].

Algorithm: General Procedure for Ensemble Method [18]

Input: dataset (set of records)

Output: classified dataset

Set: D the original training data, k the number of base classifiers and T is the test data

For i = 1 to k do Create training set, Di from D.

    Build a base classifier Ci from Di.

    End For

For each test record x ∈ T do

    C*(x) = Vote(C1(x), C2(x), ..., Ck(x))

End for

3.2 CORRELATION MEASURE

We can compute the correlation among features based on multimethods, one of those methods is called the Pearson correlation. In general, we used it to determine which of these features is more effective in taking decision or to determine the positive and negative association among features. This mechanism is based on three criteria:

- If the correlation between features goes towards (+1) this indicates a positive relationship between the features.
- If the correlation between the features goes towards (-1) this indicates an inverse relationship between the features.
- Between any two features, if the correlation is (0), then there is no correlation between them.

3.3 CONFUSION MATRIX

The Confusion Matrix is a particular arrangement of a table that lets visualization of performance of a Tree Net. Below, a table illustrate the observations that represent the testing dataset.

Table 1. Comparison of the matrix of confusion between the forecasting class and the observations

Predicted	Observed	
	Positive: True	Negative:False
Positive: True	TT	TF
Negative: False	FT	FF

Several standard terms for the two confusion matrix classes were identified as follows [17]

Accuracy (ACC): ratio of the overall number predictions that is correct is calculated as follows:

$$ACC = \frac{(TT + FF)}{(TT + TF + FT + FF)}. \qquad (2)$$

The R is a measure of the number of true cases in confusion matrix, can be computed as follows:

$$R = \frac{TT}{TT + FT}. \qquad (3)$$

Finally, precision (PT) is the proportion of the predicted positive cases that are correct, can be computed as follows:

$$PT = \frac{TT}{TT + TF}. \qquad (4)$$

The F_measure takes into account both accuracy and recall to supplying measurement for the system that is single:

$$F = \frac{2.Precision.Recall}{Precision + Recall}. \qquad (5)$$

The F measure was derived in order to F ß measures the analytical performance (Narayana,et al, 2012). The parameter ß basically takes the values 0.5, 1 and 2, with the higher values of ß laying greater emphasis on securing the analysis right in the sense of achieving a higher positive analysis value

$$F_\beta = \frac{(1 + \beta^2).(Precision.Recall)}{(\beta^2.Precision + Recall)}. \qquad (6)$$

4. SUGGESTED SYSTEM

The proposed system used raw dataset to analyse credit card databases depending on Treenet technique. This system includes the following steps; (a) find the suitable database from the internet because this database is not available in Iraq, (b) pre-process that credit card database based on Pearson correlation matrix to determine which feature is less correlated with others to remove it and reduce the time of prediction, (c) split pre-processing database into two parts: training and testing dataset, (d) apply TreeNet prediction data mining techniques (TPDMT) on training dataset to test if we need payment next month or not, also, find the optimal tree. Treenet based on Boosting Machine uses Decision Trees (DTs) usually to generate the predictor. (e) Finally, pass the testing dataset on the optimal tree results from TPDMT, then five experimental measures created by a confusion matrix would be assessed, namely: accuracy (AC), recall or true positive rate (TP), precision (P), F-measure (consideration of both precision and recall) and Fb Fig. 2 shows the structure of the proposed method.
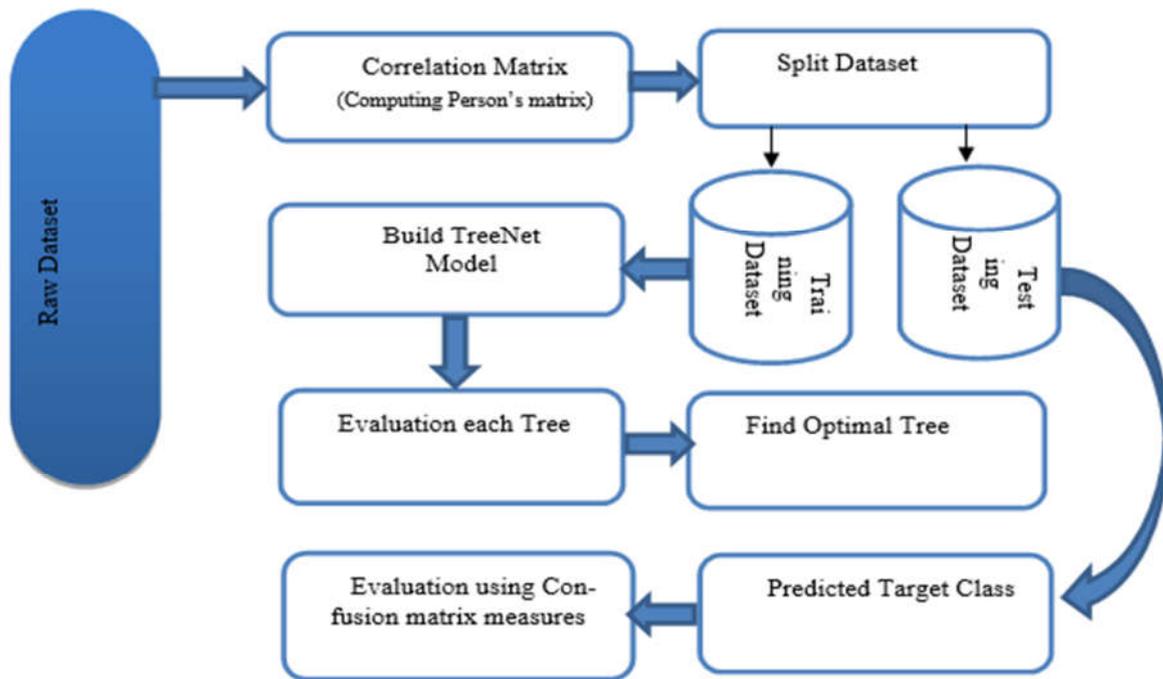

Figure 2 – Block diagram of the proposed system

5. IMPLEMENTATION

The proposed system is implemented on a raw dataset to receive the results.

Step 1: The dataset used [21]
Description of raw dataset
- Records: 1000 from total 30000
- No. of Features: 23
- Description of features:
  - X1: Amount of the given credit (NT dollar): it includes both the individual consumer credit and his/her family (supplementary) credit.
    - X2: Gender (1 = male; 2 = female).
    - X3: Education (1 = graduate school; 2 = university; 3 = high school; 4 = others).
    - X4: Marital status (1 = married; 2 = single; 3 = others).
    - X5: Age (year).

  - X6 - X11: History of past payment
  - X12-X17: Amount of bill statement

- Target called default payment next month (Y), it takes a binary variable, default payment (Yes = 1, No = 0), this means N Classes: 2

Step 2: Compute the correlation among the features using (Pearson's correlation matrix)

As a first step, we apply the correlation function from Pearson type to compute the relation among the features and remove the features that have less correlation with others, but in this work, the features have strong relation with others, therefore, we do not remove any feature. Table-2 shows the Pearson's correlation matrix of credit card database.

Table 2. Person's correlation matrix of Credit Card

From the above Table we found that all the features are important, therefore, we could remove any feature or consider it as less important than others.

Step 3: **Split the dataset to training and testing stage**

In this step, we split the database into two parts, training and testing dataset as explained in Table 3.

Table 3. Split Credit Card dataset into training and testing parts.

	N	Proportion
Learn	795	0.795
Test	205	0.205

Step 4: Apply TreeNet Data Mining Algorithm as Classification Methods

In this step, we use the TreeNet on the training dataset, the main purpose of this step is to determine the optimal tree based on multi evaluation measures; in general, the max number of trees uses 200 trees.

Step 5: Evaluation of 200 Tree based on Multi Evaluation Measures

Table 4 explains the result of evaluation of the tree based on multi measures.

Table 4. Evaluation of the trees based on multi measures

Trees	AveLL		ROC		Lift		KS-stat		Class.Error	
	Learn	Test	Learn	Test	Learn	Test	Learn	Test	Learn	Test
1	0.53461	0.45855	0.73333	0.60941	2.61896	1.78977	0.42357	0.28259	0.22767	0.16098
10	0.51951	0.44969	0.78900	0.66693	3.32873	2.42424	0.45468	0.32082	0.22767	0.16098
18	0.50848	0.44355	0.79967	0.67909	3.17680	2.72727	0.48853	0.34531	0.22767	0.16098
20	0.50573	0.44262	0.80386	0.67195	3.23204	2.72727	0.49143	0.32082	0.22767	0.16098
30	0.49354	0.43672	0.80892	0.67389	3.19751	3.03030	0.50842	0.32787	0.22767	0.16098
31	0.49269	0.43637	0.80868	0.67336	3.18370	3.18182	0.50440	0.32082	0.22767	0.16098
40	0.48313	0.43245	0.81206	0.67283	3.28729	2.87879	0.51331	0.32082	0.22767	0.16098
50	0.43759	0.42888	0.81502	0.67336	3.39779	3.03030	0.50976	0.32206	0.22767	0.16098
60	0.46599	0.42677	0.81795	0.67072	3.50829	3.03030	0.50721	0.32206	0.22767	0.16098
70	0.45872	0.42476	0.81878	0.67037	3.45304	2.72727	0.50948	0.30585	0.22767	0.16098
80	0.45294	0.42317	0.82113	0.67574	3.64641	2.57576	0.51500	0.32082	0.21509	0.17073
90	0.44738	0.42347	0.82510	0.68023	3.64641	2.72727	0.52052	0.32664	0.19623	0.16098
100	0.44212	0.42034	0.82779	0.67777	3.56354	0.42424	0.52151	0.32082	0.18868	0.15610
110	0.43698	0.42022	0.83175	0.67856	3.61878	2.72727	0.53483	0.32082	0.18239	0.15610
120	0.43182	0.41909	0.83521	0.68411	3.72928	2.57576	0.53745	0.31501	0.18113	0.15610
123	0.43232	0.41909	0.83532	0.68393	3.70166	2.57576	0.54234	0.31501	0.17987	0.15122
130	0.42761	0.41908	0.83821	0.68305	3.70166	2.57576	0.54269	0.31501	0.17987	0.15610
134	0.42584	0.41886	0.83870	0.68552	3.72928	2.57576	0.54432	0.31501	0.17736	0.15122
137	0.42426	0.41885	0.84057	0.68464	3.75691	2.57576	0.54106	0.31501	0.17736	0.15122
140	0.42323	0.41903	0.84143	0.68446	3.72928	2.57576	0.53908	0.31501	0.17736	0.15122
150	0.41920	0.42007	0.84425	0.68006	3.78453	2.57576	0.53468	0.30920	0.17610	0.15610
160	0.41493	0.42066	0.84817	0.68217	3.67403	2.57576	0.55572	0.31501	0.17358	0.16098
170	0.41052	0.42186	0.85215	0.67970	3.67403	2.57576	0.55763	0.29757	0.17107	0.16098
180	0.40656	0.42199	0.85459	0.67530	3.72928	2.72727	0.56252	0.29175	0.16478	0.16098
190	0.40280	0.42203	0.85868	0.67953	3.75691	2.72727	0.57386	0.29175	0.16226	0.16098
196	0.40028	0.42199	0.86130	0.67953	3.70166	2.72727	0.58037	0.29757	0.15975	0.16098
200	0.39890	0.42194	0.86316	0.68023	3.83978	2.72727	0.57874	0.29757	0.15723	0.16098

Step 6: Determining the best (Optimal) tree based on each measure

Table 5. Optimal tree Model based on each measure

Optimality Criterion	N-trees	Performance
AveLL	137	0.41885
ROC	134	0.68552
Lift	31	3.18182
KS-stat	18	0.34531
Class. Error	123	0.15122

Step 7: Determining the importance of each variables in best training model

Table 6. Variable Importance for the 137-tree Model

Var.Abs.	Importance	Var.Abs.	Importance
X6	100.00000	X14	29.06420
X20	43.53669	X23	28.18845
X18	38.87962	X17	27.27543
X5	35.51574	X10	26.68513
X8	35.44021	X13	23.88499
X21	34.21611	X16	20.92117
X22	34.11972	X7	19.47205
X15	33.02073	X3	10.96163
X12	31.14849	X4	8.39434
X19	29.96723	X11	7.89091
X1	29.75404	X2	4.74327
X9	29.71424		

Step 8: Pass testing dataset to optimal Model to analyze the results based on confusion matrix measures and Predicted Target Class

To obtain results, we passed testing dataset to optimal model to analyze the results based on confusion matrix measures and Predicted Target Class, we use the confusion matrix measures that are explained in section (3.3)

Table 7. Confusion Matrix of Testing Dataset

Actual	Predicted Class		Actual
Class	**0**	**1**	**Total**
0	168.00	4.00	172.00
1	27.00	6.00	33.00
Pred. Tot.	195.00	10.00	205.00
Correct	0.97674		0.18182

Table 8. Analysis of the Results of Testing Dataset based on Confusion Matrix Measures

Total records of testing stage	205
Accuracy	84.87%
True positive (TP)	86 %
Precision	97%
Recall	91.16%
F-Measure	94.58%
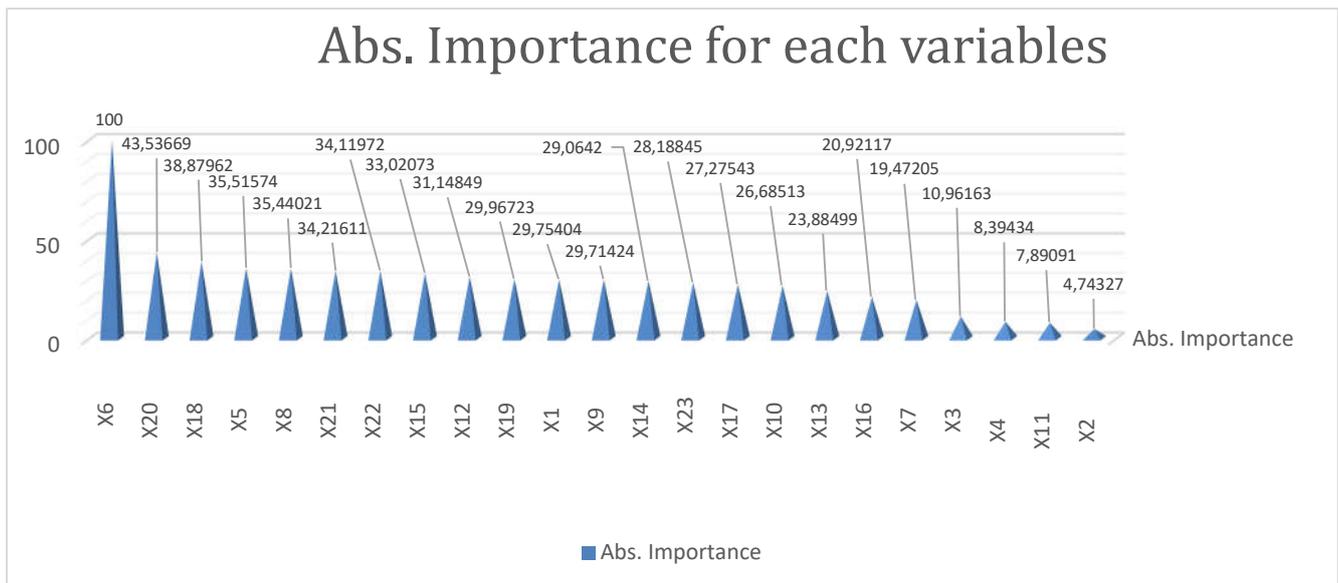
Figure 3 – Absolute importance for variables in optimal model

In general, there two variables for more contact and effect on the target in optimal predictor are (X20, X6).

6. CONCLUSION

When applying pre-processing in the first stage using Pearson correlation, it is found that the credit card features have strong correlation and we cannot

delete any of them because of their strong correlation with the target Y, the target is predicting credit card payment next month.

Using multiple decision trees representing TreeNet gives more confident results than single decision tree; and makes the optimal forward predictive model.

Although correlation has proved that all features are important and strongly correlated, but also the best decision tree achieved has proved that all features are important in decision making of data.

Using confusion matrix measures for predicting the payment possibility next month has proved its efficiency since we achieved a tree positive rate near to 0.97674, as it is shown in Table (7) and the prediction accuracy was 84.87%

7. REFERENCES

[1] Y. Cheng, C.-H. Lien, "The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients," *Expert Systems with Applications*, vol. 36, pp. 2473–2480, 2009.

[2] *Credit Card Services: Using the Simple Order API*, CyberSource Corporation HQ, July 2019.

[3] F. Provost, & R. Kohavi, "Guest editors' introduction: On applied research in machine learning," *Machine Learning*, vol. 30, issue 2, pp. 127-132, 1998.

[4] P.M. Deshpande, A.H. Siddiqi, K. Alam, K. Parmar, "Applications of data mining techniques for fraud detection in credit-debit card transactions," *Proceedings of the National Conference on Technological Advancement and Automatization in Engineering*, January 2016, pp. 339-345.

[5] R. Patidar, and L. Sharma, "Credit card fraud detection using neural network," *International Journal of Soft Computing and Engineering*, no. 1, pp. 32-38, 2011.

[6] J.K. Bae, and J. Kim, "A personal credit rating prediction model using data mining in smart ubiquitous environments," *International Journal of Distributed Sensor Networks*, vol. 11, issue 9, paper 179060, 2015.

[7] H. Modi, et al., "Fraud detection in credit card system using web mining," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 1, issue 2, pp. 175-179, 2013.

[8] I.H. Witten, et al., *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 2016.

[9] V. Bhusari, and S. Patil, "Application of hidden Markov model in credit card fraud detection," *International Journal of Distributed and Parallel Systems*, vol. 2, issue 6, pp. 203-211, 2011.

[10] M.A.Z. Khan, J.D. Pathan, and A.H.E. Ahmed, "Credit card fraud detection system using hidden Markov model and K-clustering," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, issue 2, pp. 5458-5461, 2014.

[11] H. Patel, and D. Patel, "A brief survey of data mining techniques applied to agricultural data," *International Journal of Computer Applications*, vol. 95, issue 9, pp. 6-8, 2014.

[12] *TreeNet™ An exclusive implementation of Jerome Friedman's MART methodology*, Salford Systems, 2005.

[13] *Robust Multi-Tree Technology for Data Mining, Predictive Modeling and Data Processing*, Version 2.0, 2005.

[14] D. Steinber, *Overview of TreeNet™ Technology Stochastic Gradient Boosting*, January 2009.

[15] A. Patel, S. Al-Janabi, I. Al Shourbaji, J. Pedersen, "A novel methodology towards a trusted environment in mashup web applications," *Computers & Security*, vol. 49, pp. 107-122, 2015. https://doi.org/10.1016/j.cose.2014.10.009.

[16] R. Kaviyarasi, and T. Balasubramanian, "Exploring the high potential factors that affects students' academic performance," *International Journal of Education and Management Engineering*, vol. 8, issue 6, pp. 15-23, 2018.

[17] M.J.A. Berry, and G.S. Linoff, *Data Mining Techniques: for Marketing, Sales, and Customer Relationship Management*, John Wiley & Sons, 2004.

[18] P.K. Chan, et al., "Distributed data mining in credit card fraud detection," *IEEE Intelligent Systems*, vol. 6, pp. 67-74, 1999.

[19] Y.B. Wah, and I.R. Ibrahim. "Using data mining predictive models to classify credit card applicants," *Proceedings of the 2010 6th IEEE International Conference on Advanced Information Management and Service (IMS)*, 2010, pp. 394-398.

[20] J.A. Roberts, and E. Jones, "Money attitudes, credit card use, and compulsive buying among American college students," *Journal of Consumer Affairs*, vol. 35, issue 2, pp. 213-240, 2001.

[21] https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients.

[22] G.P. Spathoulas, S.K. Katsikas, "Reducing false positives in intrusion detection systems," *Comput. Secur.*, vol. 29, pp. 35–44, 2010.

[23] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, 3th Edition, Elsevier, 2013.

[24] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*, Pearson Addison-Wesley, 2006.

Ahmed Mohammed Hussein, *a Bachelor of Computer Science in 2004 from the University of Babylon, Faculty of Science, and then got a Master's of IT from the University of Tenaga, Malaysia in 2012, and do as lecturer at the University of Babylon, College of Science for Women, Computer Department Since 2006 till now.*

Hadeel Qasem Gheni obtained a Bachelor's Degree in Computer Science from the University of Babylon, Faculty of Science for Women's, Computer Department in 2006, and then got a Master's Degree in Artificial Intelligence from the University of Babylon, Faculty of Information Technology, Software Department in 2016. Work Do as an Assistance Lecturer at the University of Babylon, College of Science for Women, Computer Department Since 2006 until now.

Wed Kadhim Oleiwi got a Bachelor Degree of Computer Science from the University of Babylon, College of Science, Department of Computer in 2005, then earned a Master's Degree from the University of Babylon, College of Science, Department of Computer in 2012 in the field of Artificial Intelligence, and work as an Lecturer at the University of Babylon, College of Science for Women, Computer Department Since 2007 till now.

Zahraa Yaseen Hasan obtained a Bachelor's Degree in Computer Science from the University of Babylon, Faculty of Science, Computer Department in 2006, and then got a Master's from the University of Babylon, Faculty of Information Technology, Software Department in 2017. Work as an Assistance Lecturer at the University of Babylon, College of Science for Women, Laser Physics Department Since 2007 until now.