E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 9, No. 3 (2022) 264–273.

Paper Type: Research Paper

Detection of Fraud in Banking Transactions Using Big Data Clustering Technique Customer Behavior Indicators

Ramez Kian[1,*], Hadeel S. Obaid[2]

[1] Nottingham Business School, Nottingham Trent University, Nottingham NG1 4FQ, UK; ramez.kian@ntu.ac.uk.
[2] College of Engineering, University of Information Technology and Communication, Iraq; hadeel.obaid@uoitc.edu.iq.

Citation:

Kian, R., & Obaid, H. S. (2022). Detection of fraud in banking transactions using big data clustering technique customer behavior indicators. *Journal of applied research on industrial engineering, 9*(3), 264-273.

Received: 02/10/2021 Reviewed: 09/11/2021 Revised: 11/12/2021 Accepted: 23/12/2021

Abstract

Human life today is intertwined with abundant trade and economic exchanges, and life would not be possible without trade and commerce. One of the main pillars of financial exchanges are banks and financial and credit institutions, which, as the vital arteries of the economy, are responsible for transferring funds and keeping the economy alive. In the world of economic competition between organizations, profitability and proper performance for stakeholders are the basic principles of the organization's survival. To increase profitability, banks must take measures that, in addition to reducing costs, increase the level of service and customer satisfaction. The best way to do this is to use new technologies and orient the bank's policies to provide services in person and independent of time and place. The use of new technologies in the banking system sometimes leads to customers' distrust and distrust of the bank. Therefore, solutions to detect fraud in banking transactions should be provided. This article aims to discover a model for face-to-face transactions and to establish a system to block fraudulently issued transactions. Therefore, a big data clustering method is designed to timely identify bribery in banking transactions. The results show that using the big data clustering method in the fastest time can detect and stop possible fraud in customers' banking transactions.

Keywords: Big data clustering, Financial transaction fraud, Fictitious transaction, Open banking.

1 | Introduction

With the development of information technology, various aspects of human life have changed. Emerging technological facilities are one of the most effective factors that have changed the nature of work of organizations, so the correct and appropriate use of this important in business is an essential factor for the survival of organizations in today's market environment [1]. Banks are no exception to this rule, but are leaders in the use of information technology tools. The shift in society from traditional banking to e-banking is widespread, for a variety of reasons, including increased productivity, reduced costs, improved distribution channels, supply and demand management, product development, and Services, increase revenue, market development, improve service quality and improve competitiveness [2].

Corresponding Author: ramez.kian@ntu.ac.uk

http://dx.doi.org/10.22105/jarie.2021.307635.1387

Today, the benefits of e-banking have become much more apparent and have an impact on economic, social, and environmental areas such as increasing internal oversight and control and combating phenomena such as money laundering, the possibility of allocating resources to economic actors, reducing crime in society, reducing pollution. The environment, and the optimal use of energy has shown [3], [4]. With all these benefits, the challenge in e-banking is security. The main concern of banks is to guarantee customers' electronic transactions, which due to their very extensive electronic banking services, it will be very difficult to detect fraud and investigate suspicious behaviors in online transactions. In the absence of security in electronic banking, in addition to financial losses, there will be greater damage, which is the deterioration of the brand and customers' trust in the bank [5]. With the increasing use of modern banking systems and increasing the number of banking transactions, financial abuses and fraud in these transactions have become more prevalent [6]. These abuses, in addition to losing the firm's financial resources, lead to a decrease in customer confidence in the use of modern banking systems, thus reducing the effectiveness of these systems in the optimal management of capital and financial transactions. Although fraud prevention is the best way to reduce bank fraud, there are ways for profiteers to achieve their goals. Therefore, methods are needed to identify suspicious transactions online and prevent them from occurring [7]. In this paper, using big data mining method, including the use of clustering algorithm, the behavior of customer transactions to investigate possible fraud in banking operations has been investigated. Many data mining algorithms and applications do not consider out-of-date data, and these data include negative percentages of accuracy and evaluation. In this article, in addition to reviewing all banking transactions, the generated out-of-pocket data is also examined. One of the best clustering algorithms is the nearest neighbor K algorithm. This algorithm considers the test sample to belong to the class with the highest number of votes among the K nearest neighbors. One of the salient features of the KNN algorithm is that, unlike other algorithms, it introduces a database or data source only once, while other algorithms examine or scan the dataset several times.

The structure of this article is as follows: in the Section 2, some of the most important facts in the field of detecting fraud in banking transactions are examined. In the Section 3, the algorithm used to detect and detect fraud is presented. Section 4 presents the results of the implementation of the algorithm and final section deals with the conclusions of the algorithm.

2 | Literature Review

Dhrawa and Patel [8] have proposed a combination of artificial intelligence, statistics and data mining techniques to detect fraudulent online banking transactions. The TRSGM algorithm consists of five main topics. The evaluation of the proposed algorithm was based on the transaction error score. The error score of this algorithm was low for transactions that are similar to normal user behavior, but for transactions that were very different from normal user behavior, it was high. Krivko [9] has introduced a new method for detecting credit card fraud in which it combines two regulated and unsupervised methods based on rules and behavioral models. In this method, the number of TP frauds detected and the number of positive PF errors were used to evaluate. Bhusari and Patil [10] offer a secret Markov model-based method for detecting credit card fraud based on modeling user behavior. The decision of this method is based on examining the possibility of a new transaction belonging to a user. In this research, it is claimed that this method can easily process a large volume of transactions compared to other methods. Randhawa et al. [11] used machine learning algorithms to detect credit card fraud. Credit card datasets available to the public have been used to evaluate the effectiveness of the model. Experimental results positively show that the proposed algorithm can well show the creditworthiness of credit cards. Mittal and Tyagi [12] used popular supervised and unsupervised machine learning algorithms to detect credit card transaction fraud in a highly unbalanced dataset. The results showed that unsupervised machine learning algorithms can detect fraud in the banking system and better classify information. Sailusha et al. [13] used two machine learning algorithms to detect credit card transaction fraud. The results of these two algorithms are accurately ranked. Comparison of the results shows that the Adaboost algorithm is highly efficient in detecting credit card transaction fraud. Rao et al. [14] presented machine learning methods to investigate credit card fraud. They collected and analyzed all the results obtained in different articles. Hussein et al. [15] proposed a

combination of multiple classifications for detecting credit card fraud based on the fuzzy method. The results obtained from the combined method in comparison with the other 7 algorithms prove that the use of the combined method can reduce the amount of fraud in banking transactions. Błaszczyński et al. [16] tested a new data set for auto loan applications using a technique not yet explored for financial fraud prediction, namely the Dominance-Based Rough Set Balanced Rule Ensemble (DRS-BRE), and after comparing it with other techniques traditionally used for predicting financial fraud, finds that the proposed approach has several advantages over the traditional ones. Zainab et al. [17] analyzed the standard models of machine learning and then boosting algorithms (ensemble technique) are analyzed to find out which perform more accurately and precisely to predict fraudulent transaction. On analyzing the outcome, it was examined that boosting algorithms gives better results. Sarma et al. [18] proposed a system to detect bank fraud using a community detection algorithm that identifies the patterns that can lead to fraud occurrences. An agile method was used to design the web-based application to detect the fraud. The application functioned as a central hub between the banks and customers. Husejinovic [19] used naive Bayes, C4.5 decision tree and bagging ensemble machine learning algorithms to predict outcome of regular and fraud transactions. Performance of algorithms is evaluated through: precision, recall, PRC area rates. Performance of machine learning algorithms PRC rates between 0,999 and 1,000 expressing that these algorithms are quite good in distinguishing binary class 0 in our dataset. Amongst all algorithms best performing PRC class 1 rate has Bagging with C4.5 decision tree as base learner with rate of 0,825. For prediction of fraud transactions with success of 92,74 % correctly predicted with C4.5 decision tree algorithm. Nayak et al. [20] are giving a machine learning model that will detect the fraud and give a known difference between fraud and genuine transactions. We use machine learning algorithms for efficient fraud detection in online transactions and represent those using graphs. The graph exhibits interdependencies between data in an effective way. Ludera [21] proposed an AIFD based on an advanced deep neural network. The network combines hybrid Synthetic Minority Oversampling and Edited Nearest Neighbours techniques that allows us to improve the detection accuracy on an European credit card data benchmark to 98%. Ananthu et al. [22] mainly focused on recognizing fraudulent transactions by analyzing the previous set of transaction records. This research work attempts to integrate big data analytics along with machine algorithms for the fast detection of large real-time data. This paper gave a brief comparison of some of the machine learning techniques using big data.

In addition to attracting customers and trying to increase customer loyalty, what is of great importance is customer retention and gaining customer trust. Provide a model for using cards to perform fraudulent transactions. Although it is not easy to detect fraud, there are various methods used to detect bank card fraud. Most of the methods used in the literature are based on data mining. Data mining methods are used as one of the main tools for detecting bank card fraud.

3 | Fraud Detection Algorithm

In this article, using macro data clustering method, fraud in customers' banking transactions has been identified. In the following, the methods used to implement the algorithm are presented. Today, data mining operations are widely used by all organizations that focus on customers. The use of big data mining helps these institutions to discover the relationship between internal and external factors. Data mining can also provide valuable services in marketing by segmenting consumers into groups with different needs and characteristics. The use of data mining in financial matters may be the first application that can be mentioned in this area. Banks have a wide range of financial services and much of the data collected in these financial institutions is complete and reliable for data mining. Good data quality facilitates data analysis and data mining operations [23].

3.1 | Data Mining System Architecture

The data mining system architecture based on *Fig. 1* consists of the following components [24]:

I. Database, analytical data warehouse, other databases that include one or a set of databases, analytical data warehouse, spreadsheets, and refining and aggregation techniques are performed on this data.
II. An analytics database or repository server that is responsible for retrieving data related to users' data mining requests.
III. Knowledge bank is a domain of knowledge that is used to guide research or evaluate the interesting results of models.
IV. Data mining engine is one of the main components of data mining systems and includes a set of data mining functions.
V. Patterns, the acquired knowledge is presented in the form of patterns and can be evaluated by their accuracy and precision functions.
VI. The user interface, as a link between the user and the data mining system, is a tool for visualizing exploration patterns in various forms.

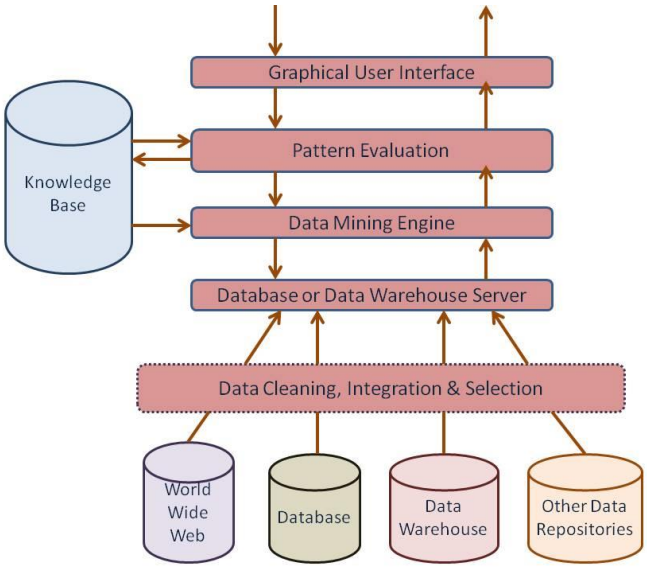

Fig. 1. Structure of data mining system.

The data mining process according to *Fig. 2* has 6 steps.

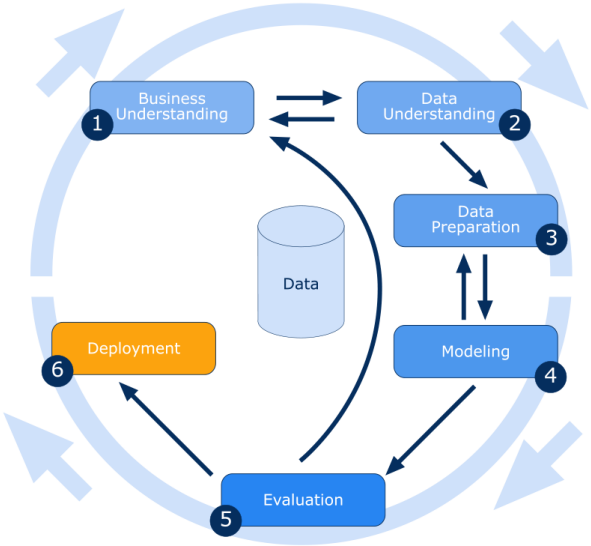

Fig. 2. Data mining process.

The Step 1 in data mining is understanding the goals of the project. At this stage we are looking at how the existing questions can be solved and how it can be brought to the data mining problems. In the field of fraud detection, the question is how to distinguish fake cases from fake ones?

The Step 2 is to understand the data and information, which includes a set of data as well as familiarity with its concept. Understanding the meanings and concepts of each variable is essential for the whole process.

The Step 3 is called data preparation, the initial collected data is structured, filtered, cleaned and transformed according to the specific goals of the project, and finally the required data mining methods are used. The selection of variables as well as other changes to the variables may also be used depending on the methods used for learning.

The Step 4 is modeling in which data mining actually takes place. Based on the prepared data, the data model is predicted or described according to the selected algorithm. This choice is very important and will affect the performance of the proposed model. For example, some classification methods can be performed with data that have a high variance, and some of them are very sensitive to outgoing data.

The Step 5 is evaluation. According to the previous step, a model for the data was presented and interpreted. It remains to be seen whether the fifth step is evaluation. According to the previous step, a model for the data was presented and interpreted. It remains to be seen whether that model answers the project's questions or just provides a well-known model. Has the ability to anticipate and help make the right business decision? In this step, we evaluate what we have in mind and the results obtained in the previous step.

The Step 6 in the data mining process is expansion or queuing. Expansion is the stage at which a comprehensive and complete model is presented. In the area of fraud detection, this means enforcing the rules created by classification methods in the marketing or business process.

Data mining is a dynamic process, so it is possible to go to the previous steps. This process should be repeated if the model does not perform well. Different data mining algorithms require data with different shapes, so data preparation is a step that may be repeated several times, or you may find that this step can never answer the questions defined in the previous steps. In this case, depending on the objectives, the data set may be modified or the number of variables may be increased [25]. Of course, some other classification methods such as after-diffusion algorithm and backup vector machines can also be used as prediction methods. In data mining, the independent variables and the dependent variable are the same properties described for each sample or observation. The values of the independent variables are usually known. However, using special methods, it is possible to predict cases in which some missing values are also found. In many cases, using non-variable conversion methods, a nonlinear problem can be solved using linear regression [26].

3.2 | Big Data Clustering

Clustering can be used in many applications. Branches such as artificial intelligence, statistics, biology, machine learning, pattern recognition, etc. can be mentioned. Although classification techniques are an effective way to identify sample groups or classes, do not forget that in such methods our samples have a class tag, so that the algorithm can design a model with its help. First, the data set is grouped with the help of similarity evaluation criteria and then each group is labeled as a class. With the help of automatic clustering, dense and non-dense regions of our data sample space are identified and also the pattern of sample distribution and correlation between the specific traits of each sample is revealed. Clustering is also used to detect out-of-range data. For example, by clustering credit card data and its transactions, it is possible to identify frequent and expensive purchases, and then to investigate the possibility of a crime to discover specific cases. In the knowledge discovery and extraction stages, clustering can also be used to preprocess data or prepare it.

3.3 | K-means Clustering

K-means clustering is a method of vector quantization, originally from signal processing that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. K-means clustering minimizes within-cluster variances (squared Euclidean distances), but not regular Euclidean distances, which would be the more difficult Weber problem: the mean optimizes squared errors, whereas only the geometric median minimizes Euclidean distances. For instance, better Euclidean solutions can be found using k-medians and k-medoids [27]. The problem is computationally difficult (NP-hard). However, efficient heuristic algorithms converge quickly to a local optimum. These are usually similar to the expectation-maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both k-means and Gaussian mixture modeling. They both use cluster centers to model the data; however, k-means clustering tends to find clusters of comparable spatial extent, while the Gaussian mixture model allows clusters to have different shapes.

The unsupervised k-means algorithm has a loose relationship to the k-nearest neighbor classifier, a popular supervised machine learning technique for classification that is often confused with k-means due to the name. Applying the 1-nearest neighbor classifier to the cluster centers obtained by k-means classifies new data into the existing clusters. This is known as nearest centroid classifier or Rocchio algorithm.

4 | Data Analysis

Considering that banks and financial institutions have found their competitive advantage in promoting and growing services in the field of electronic banking; therefore, as the popularity of customers to use electronic services increases, banks are inevitably developing in this area. What has caused banks to be outsourced their services can be the most important factor is the need for regulation. Banks are required to devote part of their capital to allocating their assets. Given that e-banking hardware must meet modern security and technology standards, it obviously imposes a heavy cost on banks. In recent years, banks have provided the infrastructure and leased bank ATMs to the private sector on the condition of ownership, and in the profits from bank transactions allocated to the bank from the use of other banks' cards on ATMs with partner banks. In this regard, some owners of private ATMs make face-to-face transactions using rental cards in order to increase the fee received. In order to discover the model/models for detecting suspected fraudulent transactions, accelerated transactions on personal ATMs have been investigated for six months. The examined information includes characteristics such as 1) card number, 2) transaction amount, 3) terminal number, 4) response code (successful/unsuccessful), 5) type of transaction, and 6) date and time of the transaction. According to the mentioned conditions, the selected statistical population has the following conditions:

- *Fraud in order to get more commission only makes sense for personal ATMs, therefore, only transactions on personal ATMs were investigated.*
- *In this data set, the focus was on transactions that were successful and received cash, because the highest fee related to receiving cash is on ATMs.*
- *Transactions related to accelerated cards (bank cards that are not related to the accepting bank) were examined.*

After clearing the data in order to make the best use of the tools available in the big data mining software, the data is aggregated based on the card number and the amount of cash transaction on a monthly basis. Using the k-means algorithm, the data are clustered in four clusters. The study of clusters indicates that the data are distributed in such a way that there is no concentration of suspicious transactions in a particular cluster. *Fig. 3* shows a graph of data distribution in clusters.

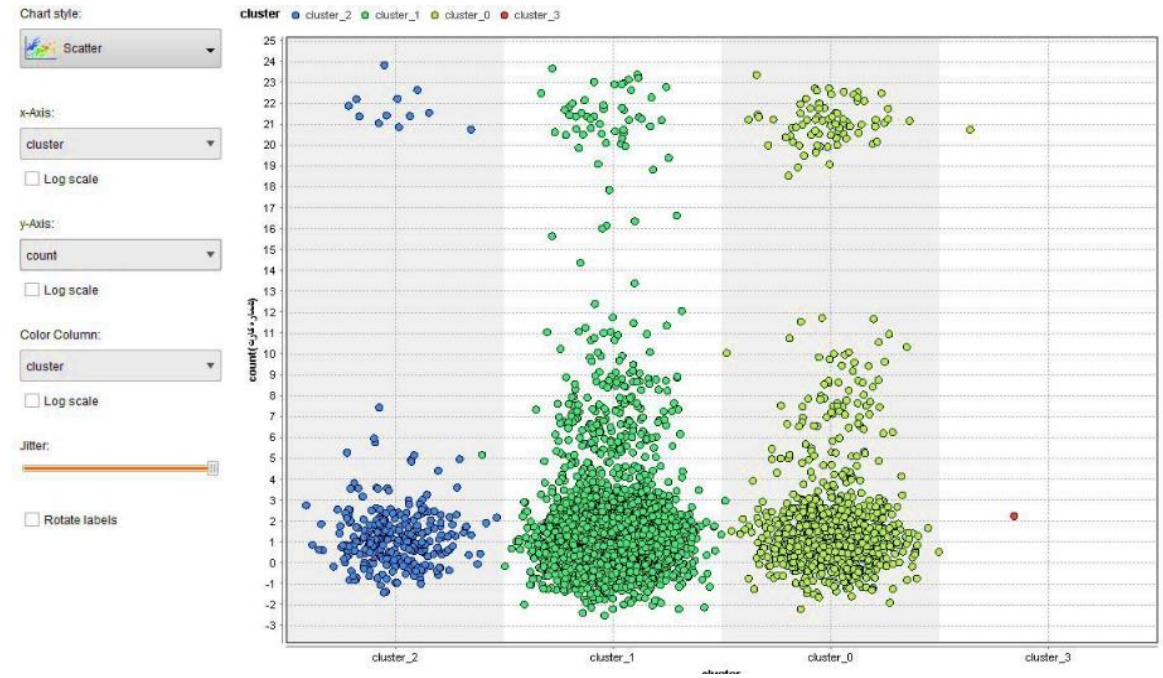

Fig. 3. Data distribution diagram in clusters.

According to *Table 1,* the average card usage and also the average transaction amount using one card during the period in different clusters do not indicate suspicious cases.

Table 1. Average amounts and frequency of card use in four clusters.

Attribute	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Transaction amount	1136846.406	737063.426	921710.526	800000
Conut	3.120	1.730	2.013	1
Sum	5336230.469	2130078.234	2923355.263	800000

After reviewing the data, it was observed that the investigation of banking transaction fraud appears in six clusters. Therefore, *Fig. 4* shows the distribution of data in six clusters.

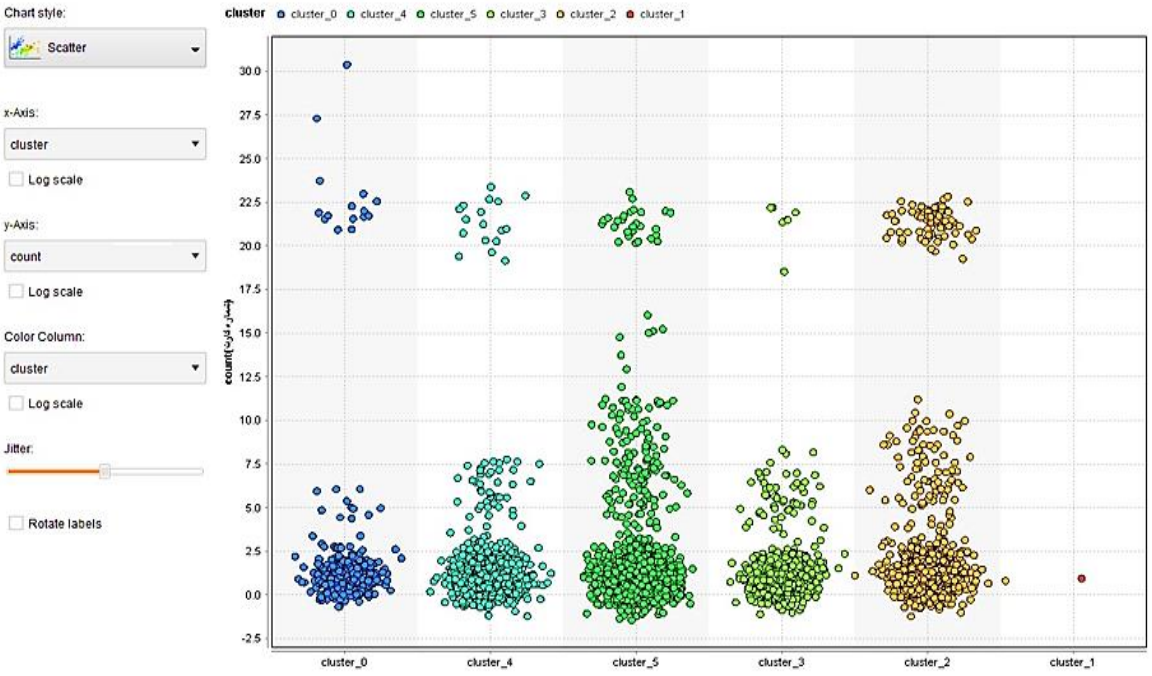

Fig. 4. Data distribution diagram in six clusters.

By examining the clusters, the concentration of suspicious cases in the bile ducts can be clearly seen *Table 2*.

Table 2. The concentration of suspicious cases in the bile ducts.

Attribute	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Transaction amount	928070.541	800000	1140294.118	792908.439	725483.871	722657.412
Conut	2.294	1	3.127	1.704	1.666	1.801
Sum	3083378.378	800000	5355686.275	2133034.111	2005752.688	2255131.383

As can be seen in *Table 2*, the highest average use of a card is in clusters of zero and two, which indicates the continued use of a card to receive cash from a specific terminal. The highest average withdrawal is related to clusters of zero and two. Considering the clustering indicators, the method used is a suitable method. Due to the suspicion of clusters zero and two, similar behaviors of these clusters are extracted as follows:

I. Continuing to use a card on a special ATM in a period.
II. Receive maximum cash from an ATM in one period.
III. Most suspicious transactions were withdrawn from ATMs between one in the morning and 4 in the morning.
IV. Other ATM services with this category of cards have not been used.
V. Suspicious transactions have been focused on one terminal and very rarely received from another terminal.

Due to the fact that the use of any technique for data mining requires validation, an innovative method has been used for validation. In this way, the cards that were in clusters other than zero and two (behavioral card numbers) have been simulated again. After simulating the data that were not present in the suspicious clusters, clustering with six clusters in *Fig. 5* was obtained again.

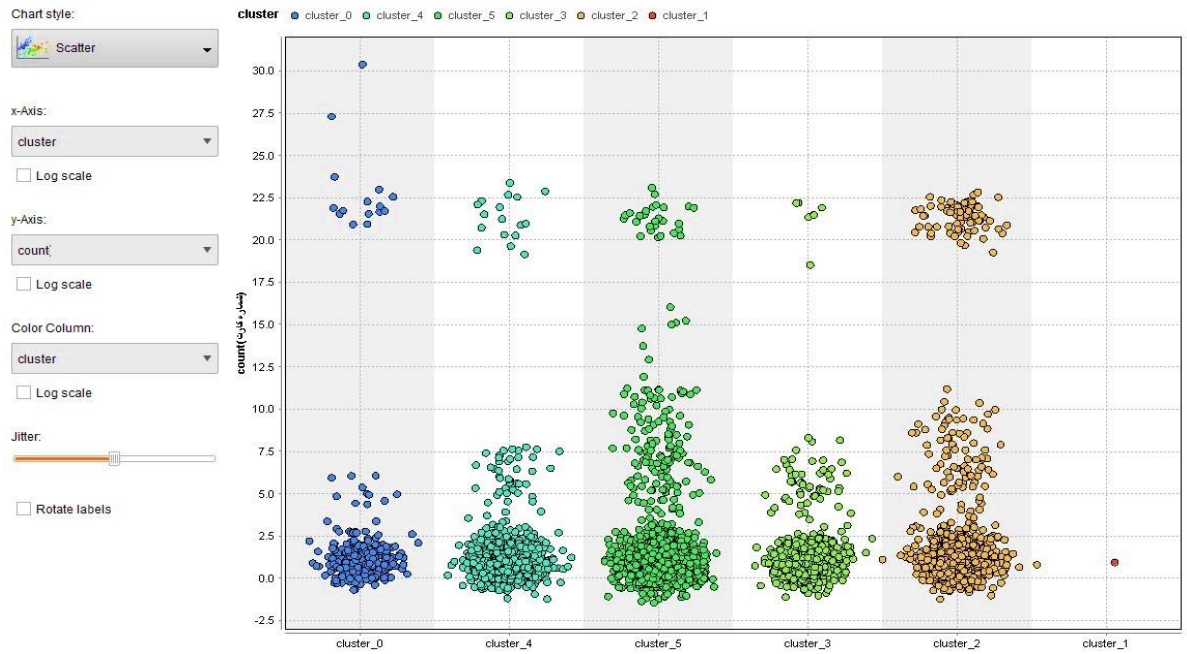

Fig. 5. Card distribution diagram after dimulation.

Based on the data performance vector, which was observed after the simulated records were placed, the Davis Boldin evaluation index considered the data clustering to be appropriate.

5| Conclusions

In recent years, various types of financial fraud, such as credit card fraud, corporate fraud and money laundering, have caused a great deal of concern and attention. Of course, the field of financial fraud detection has also witnessed significant changes. Specifically, data mining has attracted widespread attention and gained much popularity in the financial world. Successful applications of data mining have expanded in application and effectiveness. The main methods used to detect financial fraud are logistic regression models, neural networks, decision trees and clustering. Valuable information was collected by performing a simple preprocessing on the data. In fact, data scanning is done to better understand the properties of data. Due to the fact that legal requirements have caused banks to transfer their surplus assets in different sectors and to provide services through the private sector in order to develop their businesses, in this regard, it has been several years since leasing ATMs to the private sector. The ownership condition is assigned and they share the commission received from other banks with the ATM owner. In view of the above, some ATM owners, in order to increase the amount of commission received, make transactions through a number of rental cards. The purpose of this study is to identify these cards with suspicious behavior and model this behavior. First, based on the bankers' expertise, transactions related to privately owned ATMs were extracted from the data, then transactions related to accelerator cards (cards other than the accepting bank) were extracted, cleared, and clustered with a different number of clusters. Based on various indicators such as Davis Boldin, what was a good clustering was formed with k= 6. The above results were obtained in a pattern that some behaviors were similar in all suspicious cases, such as: continuing to use a card on a particular ATM in a period, receiving maximum cash from an ATM in a period. By identifying the behavior of the cards in the suspicious clusters, a model can be obtained that replaces the preventive action with the corrective action and prevents the transaction on the bank switch during the transaction with the rules extracted from the above model.

References

[1] Nouri, F. (2021). The relationship between management ability and investment opportunities with an emphasis on the role of political connection. *International journal of innovation in management, economics and social sciences*, *1*(1), 65-82.

[2] Jha, S., Guillen, M., & Westland, J. C. (2012). Employing transaction aggregation strategy to detect credit card fraud. *Expert systems with applications*, *39*(16), 12650-12657.

[3] Mohammadi, H., Ghazanfari, M., Nozari, H., & Shafiezad, O. (2015). Combining the theory of constraints with system dynamics: a general model (case study of the subsidized milk industry). *International journal of management science and engineering management*, *10*(2), 102-108.

[4] Fu, K., Cheng, D., Tu, Y., & Zhang, L. (2016). Credit card fraud detection using convolutional neural networks. *International conference on neural information processing* (pp. 483-490). Springer, Cham.

[5] Lucas, Y., Portier, P. E., Laporte, L., He-Guelton, L., Caelen, O., Granitzer, M., & Calabretto, S. (2020). Towards automated feature engineering for credit card fraud detection using multi-perspective HMMs. *Future generation computer systems*, *102*, 393-402.

[6] Tiwari, P., Mehta, S., Sakhuja, N., Gupta, I., & Singh, A. K. (2021). Hybrid method in identifying the fraud detection in the credit card. In *evolutionary computing and mobile sustainable networks* (pp. 27-35). Springer, Singapore.

[7] Hemdan, E. E. D., & Manjaiah, D. H. (2022). Anomaly credit card fraud detection using deep learning. In *deep learning in data analytics* (pp. 207-217). Springer, Cham.

[8] Dharwa, J. N., & Patel, A. R. (2011). A data mining with hybrid approach based transaction risk score generation model (TRSGM) for fraud detection of online financial transaction. *International journal of computer applications*, *16*(1), 18-25.

[9] Krivko, M. (2010). A hybrid model for plastic card fraud detection systems. *Expert systems with applications*, *37*(8), 6070-6076.

[10] Bhusari, V., & Patil, S. (2011). Application of hidden markov model in credit card fraud detection. *International journal of distributed and parallel systems*, *2*(6), 203.

[11] Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. (2018). Credit card fraud detection using AdaBoost and majority voting. *IEEE access*, *6*, 14277-14284.

[12] Mittal, S., & Tyagi, S. (2019). Performance evaluation of machine learning algorithms for credit card fraud detection. *9th international conference on cloud computing, data science & engineering (confluence)* (pp. 320-324). IEEE.

[13] Sailusha, R., Gnaneswar, V., Ramesh, R., & Rao, G. R. (2020). Credit card fraud detection using machine learning. In *2020 4th international conference on intelligent computing and control systems (ICICCS)* (pp. 1264-1270). IEEE.

[14] R Rao, H. R., Kumar, N. V., & Kolla, M. (2021). A study on machine learning approaches to detect credit card fraud. *AIP conference proceedings* (Vol. 2358, No. 1, p. 100008). AIP Publishing LLC.

[15] Hussein, A. S., Khairy, R. S., Najeeb, S. M. M., & ALRikabi, H. T. (2021). Credit card fraud detection using fuzzy rough nearest neighbor and sequential minimal optimization with logistic regression. *International journal of interactive mobile technologies*, *15*(5), 24-42.

[16] Błaszczyński, J., de Almeida Filho, A. T., Matuszyk, A., Szeląg, M., & Słowiński, R. (2021). Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. *Expert systems with applications*, *163*, 113740. https://doi.org/10.1016/j.eswa.2020.113740

[17] Zainab, K., Dhanda, N., & Abbas, Q. (2021). Analysis of various boosting algorithms used for detection of fraudulent credit card transactions. In *Information and communication technology for competitive strategies (ICTCS 2020)* (pp. 1083-1091). Springer, Singapore. https://doi.org/10.1007/978-981-16-0882-7_98

[18] Sarma, D., Alam, W., Saha, I., Alam, M. N., Alam, M. J., & Hossain, S. (2020). Bank fraud detection using community detection algorithm. *2020 second international conference on inventive research in computing applications (ICIRCA)* (pp. 642-646). IEEE.

[19] Husejinovic, A. (2020). Credit card fraud detection using naive Bayesian and c4. 5 decision tree classifiers. *Periodicals of engineering and natural sciences*, *8*(1), 1-5.

[20] Nayak, H. D., Anvitha, L., Shetty, A., D'Souza, D. J., & Abraham, M. P. (2021). Fraud detection in online transactions using machine learning approaches—a review. In *advances in artificial intelligence and data engineering* (pp. 589-599). Springer. https://doi.org/10.1007/978-981-15-3514-7_45

[21] Ludera, D. T. (2021). Credit card fraud detection by combining synthetic minority oversampling and edited nearest neighbours. In *Future of information and communication conference* (pp. 735-743). Springer, Cham. https://doi.org/10.1007/978-3-030-73103-8_52

[22] Ananthu, S., Sethumadhavan, N., & AG, H. N. (2021). Credit card fraud detection using apache spark analysis. *5th international conference on trends in electronics and informatics (ICOEI)* (pp. 998-1002). IEEE.

[23] 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.

[24] Ghahremani-Nahr, J., Nozari, H., & Sadeghi, M. E. (In Press). Investment modeling to study the performance of dynamic networks of insurance companies in Iran. *Modern research in performance evaluation*. https://www.journal-mrpe.ir/article_136608.html?lang=en

[25] Abu-Alshaikh, I. M. (2017). A new technique for investigating the dynamic response of a beam subjected to a load-moaving system. *Journal of applied research on industrial engineering*, *4*(4), 268-278.

[26] Ghahremani Nahr, J., Kian, R., & Rezazadeh, H. (2018). A modified priority-based encoding for design of a closed-loop supply chain network using a discrete league championship algorithm. *Mathematical problems in engineering*, *2018*. https://doi.org/10.1155/2018/8163927

[27] Zhao, D., Hu, X., Xiong, S., Tian, J., Xiang, J., Zhou, J., & Li, H. (2021). K-means clustering and kNN classification based on negative databases. *Applied soft computing*, *110*, 107732. https://doi.org/10.1016/j.asoc.2021.107732