Using Data Mining Techniques for Fraud Detection in The Non-banking Sector

Atef Raslan [a] , Sameh Ali *[a]

[a] Department of Software Engineer, Cario University, Egypt

**Corresponding Author: Sameh Ali [samehhussein871@gmail.com]

ABSTRACT

A method known as data mining is used to extract knowledge and insights from vast amounts of data. To find different correlational patterns several computational and statistical techniques can be used. Data mining and other tools and techniques, such as artificial intelligence, can be used to do this activity. It tries to reduce the monetary losses these kinds of operations bring and ensure the business complies with all applicable legal and regulatory obligations. To stop fraud from happening, organizations must be able to recognize it before it occurs. This paper explores the various facets of fraud detection and how businesses may use it to stop it from happening. Following that, we discuss the many methods employed in this procedure, including clustering, unsupervised learning, and neural networks. In addition, we discuss the various data pretreatment methods employed in fraud detection. These include feature selection, data normalization, and extraction. Data visualization is crucial for deciphering and comprehending mining analysis outcomes. The study then discusses the numerous data mining applications for fraud detection. These include insurance, financial statement, credit card, and healthcare fraud. We give instances of when these methods have been used to spot fraudulent activity. The limitations of data mining for fraud detection are then discussed. The significance of this technology in the battle against financial crime is highlighted in this article, which offers a thorough review of the various facets of these sectors.

1. Introduction

Unsolved deceitful practices in society's financial processes, particularly the administration of non-bank credit, call for a solution through intelligent technology [1, 2]. The accuracy and avoidance of false alarms of the existing fraud detection techniques in non-bank credit administration have not been adequate, and none have concentrated on fraud in non-bank credit default. Prediction accuracy is also impacted by deceitful duplication, missing data, and undefinable fraud scenarios. Financial fraud is any illegal act committed by people or computers that results in personal benefit at the expense of organizations or legitimately entitled individuals. However, an error must not be mistaken for fraud. Economic sabotage is the term used to describe financial wrongdoing when taken as a whole [3]. It is concerning how quickly microfinance organizations lose money that should have gone to loan recipients because of default. This trend results from the closure of numerous microfinance businesses, the denial of loans to potential clients, and the loss of numerous employments in the microfinance industry and other industries. Using data mining to foresee fraud in managing non-bank loans and avoid loan default. According to loan records from the past, a manual examination by a credit officer would not have shown this fraud. However, data mining tools can find these obscure patterns. The accuracy capabilities of statistical and conventional techniques in this direction are constrained. Case-based, analogy-based, and statistical approaches have been employed; however, they cannot identify 21st-century fraud attempts; thus, data mining tools use the method to anticipate fraud because credit history judging by humans is inefficient owing to the volume and a variety

of data. [3].

Businesses use the practice of data mining to examine and find links and trends in their data. They can enhance their decision-making and streamline their procedures with its assistance. Sadly, the proliferation of information has also contributed to increased fraud. Data mining and fraud detection have become increasingly important in today's society due to the rise in cybercrime and their importance. Data mining is extracting important knowledge and insights by analyzing vast amounts of data. Uses several tools to stop unauthorized actions for fraud detection [4].

Misusing a profit-making organization's system is called fraud in this context, even though no immediate legal repercussions may result. If fraud is rampant and the preventative measures are not foolproof, it might become a business-critical issue in a setting with fierce competition. As a component of total fraud control, fraud detection automates and reduces the manual components of a screening/ checking process. One of the most well-known industry/government data mining applications is now in this field [5]. One cannot be completely decisive regarding an application's or transaction's validity and motivation. Given the situation, the most economical course of action is to use mathematical algorithms to extract potential fraud evidence from the accessible data.

This paper examines the various facets of fraud detection and how businesses may use it to stop it from happening. It explores the applications of these techniques in various areas, such as healthcare fraud, credit card fraud, financial statement fraud, and insurance fraud. Additionally, it emphasizes the significance of data pretreatment. The paper also emphasizes the significance of fraud detection and data mining methods in today's environment. It encourages businesses to adopt these techniques in their fraud prevention strategies.

This research provides information regarding the regulation, treatment, and detection of fraud in various areas, such as healthcare fraud, credit card fraud, financial statement fraud, and insurance fraud. This contributes to fraud detection and reduces the risk of fraud in areas referred.

2. Literature Review

Data mining techniques have gained popularity as a way to uncover hidden insights from huge databases. Identifying fraudulent activity and anomalies is one of this technology's most crucial applications. These variations from the predicted trends may point to a problem. The prevention of fraud is crucial to the day-to-day operations of the banking and insurance sectors. This study examines the numerous research that investigates the use of data mining techniques in this field.

The authors introduced Almost all technical and review articles on automated fraud detection published during the past ten years are categorized, compared, and summarized in this survey work. The term "professional fraudster" is defined, the primary types and subtypes of recognized fraud are formalized, and the adversary, fraud subtypes and types, technical aspects of data, performance indicators, and methods and tactics are all defined. This study demonstrates how the "fraud detection" field might profit from other related fields after identifying the limitations in approaches and strategies. Future "fraud detection research", can benefit particularly from unsupervised methods from counterterrorism work, actual monitoring systems, and text mining from law enforcement, as well as semi-supervised and game-theoretic methods. Approaches from intrusion and "spam detection" communities. Fawcett and Provost (1999) successfully applied their "fraud detection method" to news story monitoring but unsuccessfully to intrusion fraud detection [5].

R. Gupta introduced Fraud detection as a procedure that involves spotting and stopping fraudulent activity. Data mining and other technologies and methods, including artificial intelligence, are used to carry out this operation. It tries to reduce the monetary losses these kinds of operations bring and ensure the business complies with all applicable legal and regulatory obligations. Additionally, this paper explores the numerous data mining applications for fraud detection. And how organizations may use it to stop it from

happening. It also requires a coordinated strategy with diverse tactics, including audit trials and human intervention. The significance of this technology in the battle against financial crime is highlighted in this paper, which offers a thorough review of the various facets of this sector. Following that, we discuss the many methods employed in this procedure, including clustering, unsupervised learning, and neural networks. In addition, we discuss the various data pretreatment methods employed in fraud detection. The study then discusses the numerous data mining applications for fraud detection. These include insurance, financial statement, credit card, and healthcare fraud. Data mining techniques can still offer a variety of advantages to businesses. They can assist in minimizing losses and enhancing operational effectiveness by assisting in detecting and preventing fraudulent conduct. We may anticipate the development of more advanced fraud detection systems as the approaches advance. One of the main issues prohibiting businesses from effectively exploiting fraudsters is their sophistication. The models must be updated frequently because of this. The security of the data mining techniques' collected data is one of the ethical concerns that must be considered [4].

Examples of fraud in the paper include insurance fraud, credit card fraud, and others. Therefore, detecting fraudulent behavior is essential for limiting these expenses. To detect bank fraud, this article uses data-mining techniques, for example (classification, forecasting, clustering, and association, to examine customer data and find trends that could indicate fraud. Once the patterns are found, a higher degree of verification and authentication can be introduced to the banking procedures. The scammers keep track of customers' account balances after memorizing information from their bank transactions, cash withdrawal branch transactions, and ATM transaction points and locations. A second authentication is necessary when withdrawing from a new branch or point or exceeding the usual amount range. This prevents transactions from being completed unless authentication is successful. The paper recommended that the real-time fraud detection proposal will save banks from significant losses. It will also protect the clients from suffering financial loss. In addition to improving society by reducing financial losses, it will increase people's confidence in putting their money in banks. It will deter scammers from carrying out their schemes in the future. According to the research's findings, biometrics continues to be the ideal authentication method. In risk assessment, it is not a hard-and-fast rule that two clients with the same traits or attributes would behave similarly, but in most cases (more than 65%), the behaviors will be the same [6].

This paper examines the financial effects of corporate fraud. The recent increase in corporate fraud is particularly important for investors and other market players. The study focuses on the borrowing habits of Chinese businesses following instances of corporate fraud, and we discover that these businesses' bank loans after punishment are not only much smaller but also smaller than those for businesses that did not commit fraud. In addition, We show that corporate dishonesty affects the "performance bank loan" relationship indirectly because loan interest rates are not merely higher after punishment than before. Our findings imply that corporate fraud has a detrimental impact on a company's ability to obtain debt financing, offering new proof of the economic effects of fraud. Corporate fraud affects the link between bank loans and performance, affecting bank loan policy [8].

The authors introduced the use of the support vector machine. Instead of approving loans that would normally not be granted, this work uses a supervised learning approach based on machine learning to estimate the likelihood of fraud in a loan application. As the study indicated, until Credit or loan defaults have the closure of many banks, also, many nations have experienced recession due to the people's lives being difficult and the dishonest loan administration practices. A method utilizing "SVM" in "kernel mode" produced an accuracy of 81.3%. Loan scrutiny, when applied, has good true positive and false negative rates to show its value. The Python program was used to identify 129 fraudulent transactions from the testing data. Naive Bayes, decision trees, and neural networks are additional extant methods. However, utilizing the dataset used in this work, the SVM is more accurate than the abovementioned approaches. The work has

demonstrated the SVM effectiveness in a novel setting. High accuracy is reached with a significantly lower false positive rate and a higher true positive rate [9].

This study, which centers on the Kenyan banking sector, attempts to advance knowledge and understanding of fraud in Kenya's financial institutions. The research was conducted using a theoretical framework based on the Fraud Triangle to examine the prevalence of fraud and the motivations of fraudsters. The approaches employed in the industry for detecting and preventing fraud were conventional and equivalent to those used globally. The Fraud Triangle, concluded from this study, is less successful in explaining the cooperative and predatory behavior of the Kenyan bank fraudster. Both internal and external factors have a role in fraud in Kenya. Several issues have been found, such as poor industry cooperation, poorly trained police and prosecutors, inefficient justice systems, weak government regulatory frameworks, and low. According to the survey's findings, most fraud in Kenya's banking sector is neither particularly high-volume fraud nor technologically advanced fraud. Only a few occurrences of fraud included credit cards; the bulk of frauds involve misdirection, theft, or appropriation of cash or checks. Inadequate fraud controls are an issue that plagues the global banking industry. Thus, the Kenyan banking industry is probably not very different regarding experienced fraud than the global average. However, some differences are created by the industry's technology level. Most fraud is investigated internally due to fear of damaging the banks [10].

This study thoroughly reviews the state-of-the-art research in classifying financial fraud based on forms of fraud and data mining from 2009 to 2019 inclusive. It was confirmed in the study that technologies are used in detecting financial fraud. The review's findings produced a sample of 75 pertinent articles divided into four major categories (bank fraud, insurance fraud, financial statement fraud, and cryptocurrency fraud). They included 58 conference papers and 17 peer-reviewed journal articles. And, the study demonstrates the usage of 34 data mining techniques to spot fraud across different financial apps. The study proved that one of the most popular methods for spotting financial fraud is the SVM, which accounts for around 23% of the study's total. Naive Bayes and Random Forest both come in second with 15% each. Our thorough evaluation showed that 61 out of 75 research studies—or 81.33% of the total number of papers—showed that most data mining techniques were extensively used to combat insurance fraud and bank fraud. With relevant information on the most important data mining techniques and a list of nations vulnerable to financial fraud, this review is a good reference tool for assisting academic and practical industries in identifying financial fraud. Our evaluation contributes by increasing the sample of the articles excluded from other studies and by summarizing the important work produced by numerous academics in financial fraud [11].

This study aims to use data mining to build a model that can find hidden patterns and detect and forecast fraud based on those patterns. Two objectives guided the study: firstly, to find hidden patterns and trends in data, and secondly, to construct a model that predicts and forecasts fraud based on discovered patterns. The patterns of odd transactions with an upward tendency were found using the suggested model. More research into data mining for fraud detection is crucial, as the prediction score and forecasts suggested that these transaction patterns were likely to persist and present significant risks of fraud occurring. The study developed a data mining model for fraud prediction and forecasting. Been employed in data mining tasks such as clustering and time series analysis from which we could get the patterns and trends. Some transactions may be legitimate and others not; therefore, depending on the severity and business rules of the bank, sound decisions should be made and action taken accordingly. "The study recommended" that this model be implemented and adopted to predict and forecast fraud patterns for better knowledge of the types of transactions. The proposed model will provide actionable insight and profitable results by uncovering hidden patterns of transactions and help alleviate the chances of fraud [12].

This paper presents the application of intelligent techniques using data mining techniques such as

Multilayer Feed Forward Neural Networks (MLFF), Genetic Programming (GP), and Support Vector Machines (SVM), To identify businesses that commit financial statement fraud, Logistic Regression (LR), the (GMDH) Group Method of Data Handling, and the Probabilistic Neural Network (PNN) were used. A dataset with 202 Chinese is used to test each of these methods. Enterprises comprise the dataset, examined using stand-alone algorithms such as MLFF, SVM, GMDH, GP, LR, and PNN. The top 18 features are then chosen in the first case using the t-statistic, while the top 10 features are chosen in the second scenario. The classifiers MLFF, SVM, GMDH, GP, LR, and PNN are once more used with the decreased feature subset. Results based on AUC showed that the PNN performed best, followed by GP, which generally produced slightly lower accuracies. This study's findings are also superior to those of an earlier investigation into the same topic [13].

The literature analysis indicates that data mining techniques are increasingly used in various sectors, including healthcare, banking, and non-banking, to identify fraud and abnormalities. These strategies can aid in locating prospective con artists and suspicious activity that would go undiscovered using more conventional tactics. It has been demonstrated that several data mining methods, including neural networks and association rule mining, are efficient at spotting fraud. Additional study is required to create effective and precise approaches for this sector. The many facets of data mining techniques and how they are used in fraud detection require further research. They are anticipated to increase system accuracy and lower fraud-related financial losses.

3. Fraud detection in the non-banking sector
In the Non-banking sector, data mining may be used to detect fraud. Since Detecting fraud is a priority for many businesses, data mining has increased the amount of fraud being identified and reported. Financial organizations have developed two separate methods to identify fraud trends. The microfinance company can then check for indications of internal issues by comparing those patterns to its database. The Non-banking industry is putting in more effort to detect fraud. Fraud management requires a lot of expertise. It is crucial in detecting fraud since it indicates which transactions the user did not approve[17]. One essential use of data mining in the finance industry is fraud detection. Too many businesses are concerned about being able to spot fraudulent conduct, and data analysis is helping to find and report more suspicious transactions. Financial institutions have created two distinct methods for identifying fraud trends.

4. Data Mining Model
The descriptive model highlights the patterns or relationships in the data and then investigates what constitutes its studied data [6]. The data mining models and tasks are displayed in Figure 1.

5. Data Mining Tasks
Depending on how the data mining findings will be used, various sorts of data mining tasks exist. These data mining jobs fall under the following categories [7].
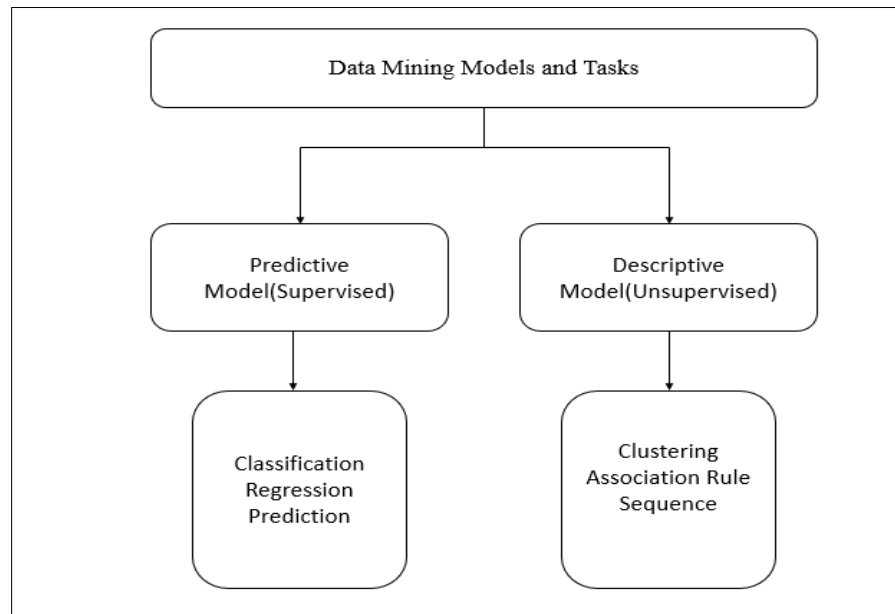
Figure 1 Data Mining Models and Tasks.

- Exploratory Data Analysis: is the act of merely examining the data without having a specific goal in mind. These methods are visible and interactive.
- Descriptive Modelling: It includes models for the overall probability distribution of the data, grouping the p-dimensional space, and models defining the connections between the variables. It describes all the data.
- Predictive Modelling: With this model, the value of one variable can be predicted based on the knowledge of the values of other variables.
- Discovering Patterns and Rules: By identifying parts of the space describing the various sorts of transactions where the data points are considerably different from the rest, pattern detection techniques are used to identify fraudulent behavior.
- Retrieval by Content: It is locating Indicators in the data collection that resemble the patterns of interest. Most of the time, text and image data sets are used for this activity.

6. Overcoming Losses to Fraud
According to a 2011 fraud survey by FICO, the following are necessary to reduce fraud losses significantly [6].
- Fraud detection in real time.
- Analytics
- Workflow
- Efficient rules engine

7. The main steps in using data mining to detect fraud.

Businesses and financial institutions lose millions of dollars annually because of fraud. Identity theft, money laundering, and credit card fraud are examples. Simple fraud detection techniques, including manual reviews, cannot catch complex fraud schemes. Many organizations now use data mining tools to spot and prevent fraud because of the rise in fraud cases. This method entails analyzing the data gathered from numerous sources to find anomalies and trends. As shown in the figure.2
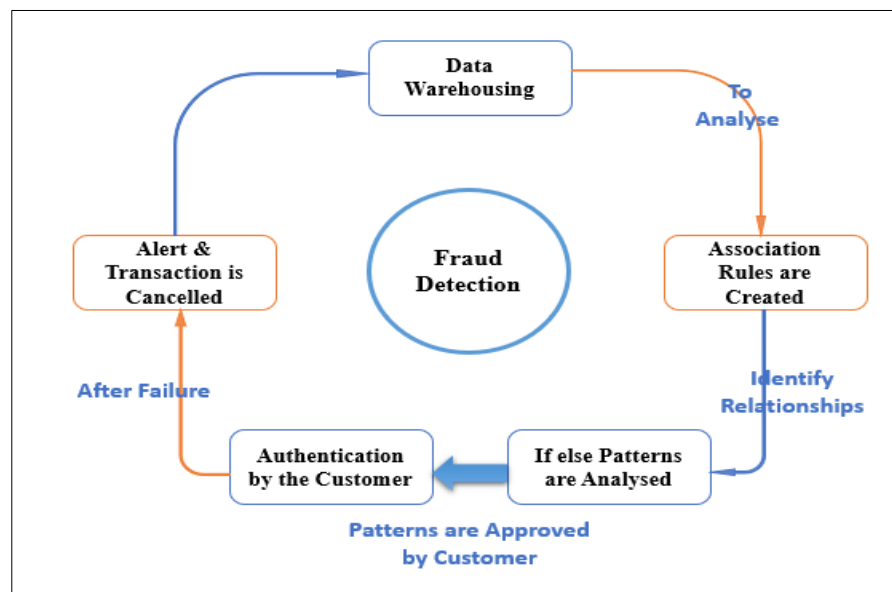
Figure 2 Steps involved in identifying fraud.

- Data Collection: The initial action in data mining to detect fraud is gathering the appropriate data from diverse sources. Databases, transaction histories, user profiles, and logs can all be used for this. The information gathered should contain all pertinent information on the transaction, such as the time and date of the transaction. Additionally, gathering as much data as possible is critical to have a full picture of the consumers and transactions.
- Data Preprocessing: The preprocessing phase entails completing missing data, eliminating duplicates, and addressing inconsistencies before employing data mining algorithms. Additionally, this phase aids in preparing the data for algorithmic usage. The missing data are filled up using various methods, including "K-nearest neighbors" and mode imputation. The choice of characteristics is crucial for detecting fraud.
- Data Mining: Techniques are employed to find probable fraud tendencies. These consist of categorization, association rule mining, and clustering. These methods assist in identifying related transactions and categorizing them as fraudulent or not. The multiple transactions in a particular location are identified using the association rule mining technique. After that, the data is analyzed using a mix of methodologies and training models.
- Fraud detection: This step is completed once the data gathered from multiple sources, including credit card transactions and bank statements, has been thoroughly examined. This phase entails analyzing the results and conducting extra research to ascertain whether the activity was fraudulent.

- Fraud Prevention: The following step is implementing efficient fraud prevention mechanisms after analyzing the data gathered from multiple sources, including social media and financial activities. These include enhancing the financial institution's security procedures, educating the workforce, and implementing fraud prevention systems. Continuous monitoring and process improvement are needed to detect fraud more accurately.

8. Methods used to identify fraud detection by data mining techniques.

Large businesses like banks, microfinance institutions, and insurance firms always seek methods to enhance their fraud detection procedures. Big data has made it more difficult to detect fraud, but it has also given researchers new chances to find unnoticed patterns and trends. They can find possible scams through data mining. By examining and spotting patterns and trends, data mining is a method that can assist businesses in identifying fraud. There are several approaches to using data mining to uncover techniques to detect fraud [4]. We presume that fraud detection has the following goals:

- To eliminate fraud to the lowest level.
- To increase customers' confidence in the non-banking and banking systems, especially for online transactions.
- To deter fraudsters (current and potential ones)

There are several data mining approaches, most of which have been applied in data mining studies. Classification, clustering, Association Rule, prediction, and sequential patterns are a few of the developed strategies.

8.1. Classification
The most widely used data mining technique is classification, which uses a set of previously categorized instances to create a model that can classify most records. Applications for assessing credit risk and detecting fraud are especially well suited for this study. Learning and classification are involved in the data classification process [6].

8.2. Clustering
Clustering is assembling related pieces of data according to their characteristics. By doing so, fraud can be avoided, and patterns and trends in the data can be identified. It can be used, for instance, to examine credit card transactions that occur in unusual settings or at odd hours.

8.3. Association Rule
Association rule mining: A technique known as association rule mining involves identifying patterns in data that often occur together. This can assist in locating potential fraud by looking for signs of behavior related to the issue. For instance, it can help identify individuals who tend to make large purchases before they get into trouble with their credit cards [4].

8.4. Prediction
Prediction is one of the strategies for data mining that identifies connections between independent and dependent variables. For instance, by utilizing prediction analytic tools, fraud can be anticipated in the banking sector. Money is an uncontrolled variable, whereas a human (the fraudster) is a dependent variable. We may then use historical data to design a regression curve to try and predict fraud. Regression analysis can be used to model the connection between at least one independent variable and dependent variables. Independent variables in data mining are already known traits, whereas responsive are the variables we want to predict. Unfortunately, many problems in the actual world are difficult to predict [6]. Several regression techniques include as following:
  - Linear Regression

- Multivariate Linear Regression
- Nonlinear Regression
- Multivariate Nonlinear Regression

8.5. Sequential patterns

One data mining technique, "sequential patterns" analysis, looks for recurring patterns in datatransactions over a certain period. Additional business analysis is conducted using the discovered patterns to identify correlations between data.

8.6. Anomaly detection

An "anomaly detection" To find anomalous data points in a huge volume of data, anomaly detection,a fraud detection tool, is typically used. It can assist in spotting probable fraud on credit cards and identity theft.

8.7. Neural networks

An algorithm for machine learning known as a neural network can identify patterns in data. It can then be used to identify fraud by analyzing the data and identifying the signs of activity that are most likelyto be fraudulent. For instance, training a neural network on a transaction credit card can identify patterns of activity that are most likely to befraudulent. Data mining is frequently used to find fraud across numerous businesses. Through the analysis of enormous amounts of data, it can assist businesses in identifying potential fraud. Clustering, anomaly detection, and classification are some methods employed in this field [4].

8.8. Rule-based systems

These systems look for probable fraud using a set of rules or heuristics. For instance, a system based on rules might highlight any transaction that exceeds a specific threshold or is made to an overseas account [14].

8.9. Decision trees

Decision trees are machine-learning algorithms that can be used to identify patterns in data. Decision trees can identify illegitimate and legal transactions [15].

8.10. Support vector machines

Are a class of machine learning methods that are useful for finding patterns in data. Support vector machines are capable of distinguishing between fraudulent and honest transactions.

9. Impact of Data Mining on Fraud Detection

Data mining can be used to find fraud patterns that were previously undetected. Machine learning techniques, which analyze previous data and identify potential fraud, can be used to accomplish this. Big data mining is crucial to the non-banking sector's fraud detection process. It can assess a lot of information data and spot behavior patterns, which can aid in averting financial loss. Data mining will probably become more important in the battle against fraud due to the development of the financial sector [4]. Data mining, used to uncover the strange behavior of data, is thought to be an efficient tool for detecting fraud. Data mining is useful for identifying fraudulent business practices and effectively protecting assets. Also, data analytics is used to identify anomalies; if those are verified, an audit inquiry can help determine whether a particular transaction is fraudulent. Furthermore, specific actions against criminals should be required to be conducted if fraudulent acts are discovered using this technique [16]
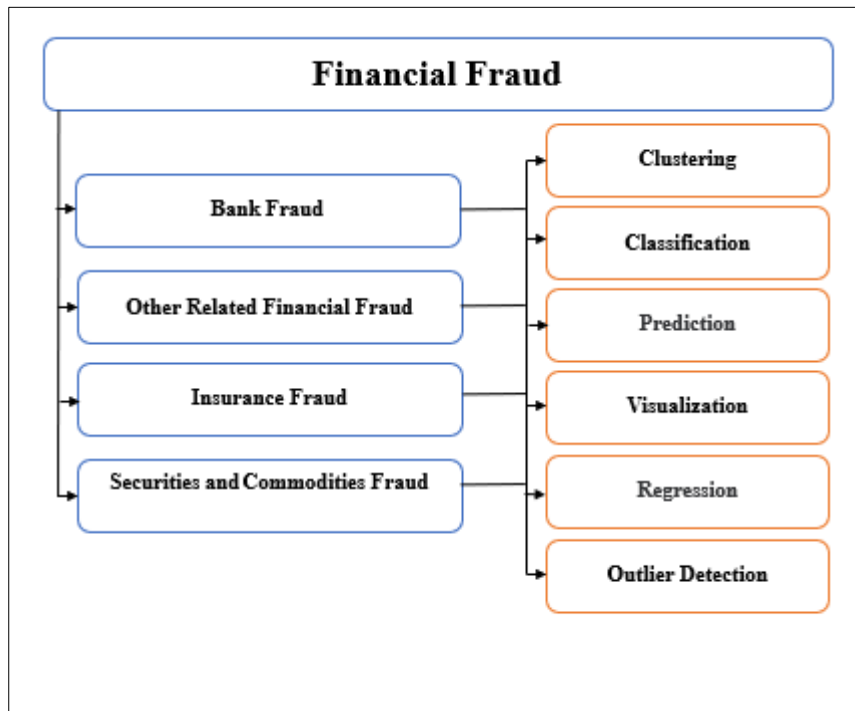
Figure 3 Importance of data mining in fraud detection.

Illustrates the main benefit of adopting data mining techniques. Additionally, several techniques, including outlier identification, classification, visualization, regression, prediction, and clustering, are used in the application of data mining processes in financial fraud. The most popular methods for spotting fraud are decision trees and neural networks.

10. Conclusion

Data mining techniques are becoming more prevalent in the fight against fraud, especially in healthcare, e-commerce, banking, and non-banking sectors. They can help organizations Determine and stop fraudulent activity. The paper provides a summary of the various techniques that are used when mining data for fraud detection. These include clustering, association rule mining, and classification. Despite the benefits of data mining methods, they still face some challenges when detecting fraud. One of the main factors that prevents organizations from using them effectively is the sophistication of the fraudsters. This is why the models must be updated frequently. One of the moral concerns that data mining techniques have to consider is the security of the information they collect. This is because the confidentiality and privacy of the data must be maintained. Despite the challenges, data mining techniques can still benefit organizations. They can help them identify and prevent fraudulent activity, which can help them minimize their losses and improve their operational efficiency. As the techniques continue improving, we expect more sophisticated fraud detection systems to emerge.

References

[1] A. A. Hameed, B. Karlik, and M. S. Salman, "Back-propagation algorithm with variable adaptive momentum," Knowledge-Based Syst., vol. 114, pp. 79–87, 2016, doi: 10.1016/j.knosys.2016.10.001.

[2] R. Koralage, "Data Mining Techniques for Credit Card Fraud Detection," Sustain. Vital Technol. Eng. Informatics, no. 2015, pp. 1–9, 2019.

[3] I. O. Eweoya, A. A. Adebiyi, A. A. Azeta, and A. E. Azeta, "Fraud prediction in bank loan administration using a decision tree," J. Phys. Conf. Ser., vol. 1299, no. 1, 2019, doi: 10.1088/1742-6596/1299/1/012037.

[4] R. Gupta, "Data Mining for Fraud Detection: An Overview of Techniques and Applications," vol. 10, no. 01, pp. 561–567, 2019, doi: https://doi.org/10.17762/turcomat.v10i1.13549.

[5]   K. S. & R. G. CLIFTON PHUA1*, VINCENT LEE1, "A comprehensive survey of data mining-based accounting-fraud detection research," 2010 Int. Conf. Intell. Comput. Technol. Autom. ICICTA 2010, vol. 1, pp. 1– 14, 2010, doi: 10.1109/ICICTA.2010.831.

[6]   S. N. John, O. K. O, and C. G. Kennedy, "REALTIME FRAUD DETECTION IN THE BANKING SECTOR USING DATA MINING TECHNIQUES/ALGORITHM," no. December, pp.   1186–1191,   2016,   doi: 10.1109/CSCI.2016.223.

[7]   A. Bhardwaj, A. Sharma, and V. K. Shrivastava, "Data Mining Techniques and Their Implementation in Blood Bank Sector – A Review," Int. J. Eng. Res. Appl., vol. 2, no. August, pp. 1303–1309, 2012.

[8]   Y. Chen, S. Zhu, and Y. Wang, "Corporate fraud and bank loans: Evidence from China," China J. Account. Res., vol. 4, no. 3, pp. 155– 165, 2011, doi: 10.1016/j.cjar.2011.07.001.

[9]   I. O. Eweoya, A. A. Adebiyi, A. A. Azeta, and O. Amosu, "Fraud prediction in loan default using support vector machine," J. Phys. Conf. Ser., vol. 1299, no. 1, 2019, doi: 10.1088/1742- 6596/1299/1/012039.

[10]   Akelola, "Fraud in the banking industry: A case study of Kenya," A PhD thesis Pap. Nottingham Trent Univ., no. July, pp. 1–422, 2012, [Online]. Available: https://core.ac.uk/download/pdf/30624246.pdf.

[11]   K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," Comput. Sci. Rev., vol. 40, 2021, doi: 10.1016/j.cosrev.2021.100402.

[12]   K. Kamusweke, M. Nyirenda, and M. Kabemba, "A Data Mining Model for Predicting and Forecasting Fraud in Banks," Proc. Int. Conf. ICT, no. November, pp. 172–177, 2019, [Online].   Available: https://www.researchgate.net/publication/33773 1727.

[13]   P. Ravisankar, V. Ravi, G. Raghava Rao, and I. Bose, "Detection of financial statement fraud and feature selection using data mining techniques," Decis. Support Syst., vol. 50, no. 2, pp.   491–500, 2011,   doi: 10.1016/j.dss.2010.11.006.

[14]   K. K. Joy Asuni, "Application of Data Mining Techniques in the Banking Sector," vol. 4, no. 1, pp.   1–8,   2022,   doi: 10.13140/RG.2.2.21091.43042.

[15]   D. Zakirov, "Application of Data Mining in the Banking Sector," vol. 4, no. 1, pp. 13–16, 2015.

[16]   S. R. Krishna, "Machine Learning based Data Mining for Detection of Credit Card Frauds," 2023 Int. Conf. Inven. Comput. Technol., no. June, pp. 72–77, 2023, doi: 10.1109/ICICT57646.2023.10134015.

[17]   F. Sabry Esmail, F. Kamal Alsheref, and A. Elsayed Aboutabl, "Review of Loan Fraud Detection Process in the Banking Sector Using Data Mining Techniques," Int. J. Electr. Comput. Eng. Syst., vol. 14, no. 2, pp. 229–239, 2023, doi: 10.32985/ijeces.14.2.12.