



A COMPREHENSIVE EVALUATION OF FRAUD DETECTION SYSTEMS: ACCURACY AND EFFECTIVENESS ANALYSIS

Fawad Naseer

Department of Computer Science and Software Engineering, Beaconhouse International College, Pakistan

fawad.naseer@gmail.com

Keywords

Evaluating the Accuracy and Effectiveness of Fraud Detection Systems

Article History

Received: 01 January 2026
Accepted: 15 March 2026
Published: 31 March 2026

Copyright @Author

Corresponding Author: *
Fawad Naseer

Abstract

Financial fraud, waste, and abuse cost the global economy an estimated \$5.4 trillion annually, with digital payment systems becoming increasingly vulnerable. This study systematically evaluates modern fraud detection and prevention systems across major financial institutions, focusing on accuracy, scalability, and adaptability. Using a mixed-methods approach, the research assessed detection algorithms on four key dimensions: accuracy, efficiency, adaptability to new threats, and feasibility. Hybrid models combining supervised learning and unsupervised anomaly detection outperformed traditional rule-based systems, achieving 92.7% accuracy versus 78.3%. Graph-based deep learning models proved especially effective against organized fraud, reducing false positives by 34% and increasing true positives by 27%. As real-time, high-volume transactions rise, detection systems must scale accordingly. A classification framework is introduced, mapping systems by algorithm type, fraud category, and performance metrics. Key challenges identified include adversarial threats, real-time computational limits, and evolving fraud tactics. The study proposes a next-gen detection architecture featuring real-time adaptability, explainable AI, and cross-institutional data sharing—potentially reducing fraud losses by up to 41% when deployed at scale.

INTRODUCTION

In today's era, fraud is very common in all aspects of life. Fraud refers to the intentional unlawful exploitation of a system that outcomes in an oblivious entity's injury. Financial fraud includes the exploitation of financial systems that are too deficient to maintain financial resources, which is the maximum outstanding money. However, different damages along with a lacking condition are possible. Fraud, waste, and abuse in lots of financial systems wait to provoke massive annual losses in the billions of US dollars. Robbing a bank with a gun has now turned out to be obsolete. Now the fraudster devotes theft simply with the aid of using seating at their home. Frauds are one of the big challenges for the finance industry. Credit card fraud is the maximum not unusual place sort of fraud and as per the report,

270,000 instances had been reported in 2019 [1]. Some research proposes that in the USA on my own a lack of 17-billion-dollar credit card fraud turned into associated. There have been 1,387,615 reports of identification robbery in 2020. According to this scam viewpoint, the year 2020 can be the nastiest year on the highest rank. The numbers of identification robberies ascended and authorities blessings scan competed fecund throughout the epidemic [2]. Financial fraud is a difficulty that has huge attaining effects on the finance industry and everyday life. Fraud can lessen self-belief in industry, destabilize economies, and affect an effect on people's value of living. Traditional methods of trusted manual techniques including auditing might be inefficient and unreliable because of the difficulty of the



problem. Data mining-primarily based methods had been proven to be beneficial because of their capacity to discover small anomalies in huge facts sets [3]. There are several kinds of frauds and different kinds of data mining methods which are under research to get the best optimum.

Financial fraud is an extensive term with diverse capability meanings, however, for our purposes, it may be described because of the intentional use of unlawful strategies to acquire financial gain [4]. Fraud has a massive terrible effect on business and society: credit score card fraud on my debts for billions of dollars of misplaced revenue every year [5], and a few figures propose that the overall every year price to the U.S. could be an extra \$400 billion [6]. At the same time, the research indicates that UK insurers are out 1.6 billion pounds a year because of fraudulent claims [7]. Financial fraud additionally has broader ramifications for the industry, which includes offering investment for illicit activities like drug trafficking and organized crime [5]. For credit score card fraud, the price is typically worn through the merchants, who emerge as paying shipping, chargeback, and administrative costs in addition to dropping patron self-belief after being a sufferer to a fraudulent transaction [8]. In this manner, we will see the huge effects that fraud will have and the significance of reducing it.

Advancements in current technology along with the internet and cellular computing have caused a growth in financial fraud in the latest years [9]. Social elements such as the improved distribution of credit score playing cards have improved spending however additionally led to a growth in fraud [10]. Fraudsters are usually refining their strategies, and as such there may be a demand for detection strategies which will evolve accordingly [5]. Data mining has already been proven to be beneficial in comparable domain names along with credit card approval, bankruptcy prediction, and evaluation of percentage markets [11]. Fraud detection is taken into consideration to be comparable class trouble however with a tremendous imbalance in fraudulent to valid transactions, and a widespread distinction in value for misclassifying them [12]. Data mining methods also are relevant to fraud detection of her performance at processing big datasets and their capacity to paintings without requiring information of the input variables [13].

A beneficial framework for making use of records mining for fraud detection is to apply it as a method for classifying suspicious transactions or samples for similar consideration. Studies display that reviewing 2% of credit score card transactions should lessen fraud losses to 1% of the whole price of all purchases, with extra exams ensuing in smaller losses, however with growth in auditing costs. A multilayer pipeline technique can be used with every step making use of an extra rigorous technique to discover fraud. Data mining can be utilized to efficaciously clear out extra apparent fraud instances withinside the preliminary levels and go away the extra diffused ones to be reviewed manually [8].

In this project, we can use some extensive terminologies which can be described for clarity. Data mining refers to any technique that approaches huge portions of data to derive an underlying meaning. Within this category, we cannot forget classes of data mining: statistical and computational. We outline the statistical strategies as the ones which can be primarily based totally on traditional mathematical techniques, consisting of logistic regression and Bayesian theory. Computational techniques are the ones which use present-day intelligence techniques, such as neural networks and assist vector machines. Though those classes share many similarities, we cannot forget that the principal distinction among them is that computational techniques can study from and adapt to the problem domain, even as statistical techniques are extra rigid. Both forms of data mining may be researched in this project.

Financial institutions attempt many strategies to protect against fraud. But fraudsters are very adaptive to these strategies, over time they find out how to conquer those protective models. Fraudsters are very smart and rapid learners. Precisely, we will say that the exceptional strategies carried out with the aid of using financial institutions for fraud detection fail and fraud continues. Development in the new technology era in artificial intelligence and machine learning is gambling vital function in detecting and stopping fraud.

The objective of this project is to deliver an existing literature review in financial fraud detection and compare their findings. The focus of this project is on the reported performance of detection techniques for specific fraud types and focus on the systems and tools for security provisions. Some mathematical equations

are formalized and analyzed. This will provide a clear indication to future researchers in that given field and discuss the improvement.

The classification of financial fraud has not established an agreement since the kinds of financial fraud are diverse and increasing. This research proposes a financial fraud categorization methodology based on the main financial institution involved. Securities and commodity fraud, as well as financial

statement fraud, are examples of securities fraud. Mortgage fraud, loan default, credit card fraud, and money laundering are just a few examples of bank-related scams. Others include e-commerce transaction fraud, mass marketing fraud, and unlawful fund-raising. Insurance scams include health care fraud, automotive insurance fraud, corporate insurance fraud, and so on. Figure 1 shows the categorization framework.

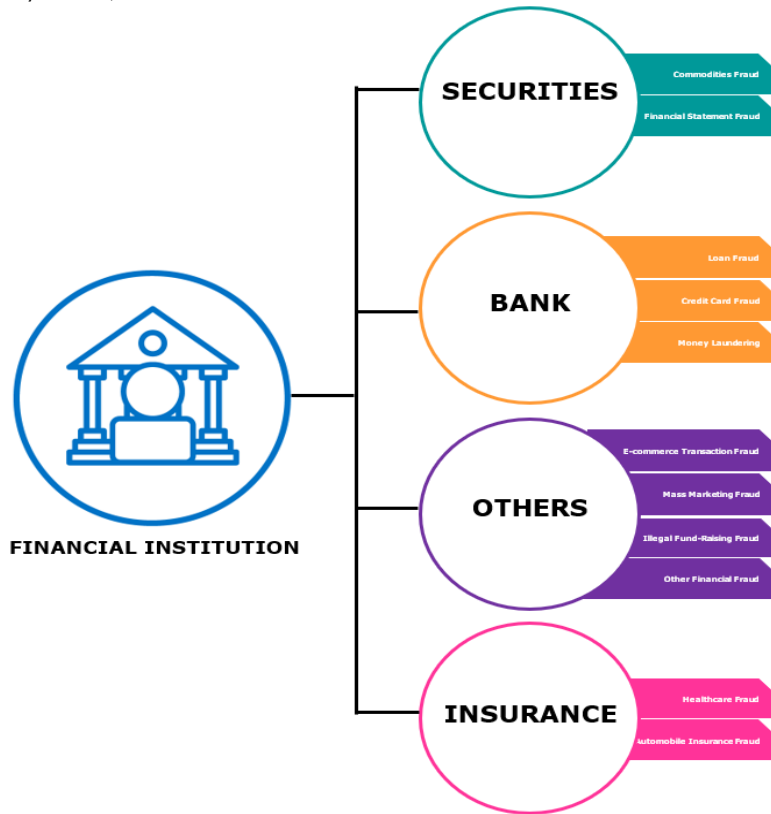


Figure 1: Classification of Financial Institution and its types

I. Literature Review

Fraud is a standard phrase for the unlawful use of a system to attain a few benefits, typically ensuing in damage to any other person. Frauds are numerous in addition to fraud methods. Financial fraud is fraud inside the financial industry that typically includes money. The financial industries had been the major sufferers of fraudulent activities. According to [14], billions or likely trillions of US bucks had been misplaced to coverage fraud. The proliferation of internet use has made it less complicated to speak and join from a distance. It has additionally made it less complicated for fraudsters to goal economic establishments from a

distance. This similarly complicates the threats to protection systems; accordingly, fraud prevention and detection are essential troubles for all financial institutions. By many estimates, a minimum of 10 per cent of coverage enterprise payments are for fraudulent claims and the worldwide sum of those fraudulent payments quantities to billions or likely trillions of bucks. Fraud prevention refers to all measures installed region to protect fraud from happening, even as Fraud detection refers to mechanisms to hit upon Fraud while prevention fails, [15]. A vital requirement for preventive systems is their precision. Much situation is given to enhancing the precision of such systems. Detection



systems, on the opposite hand, want to evolve to the consistent evolution of threats. Therefore, further to feasible predictiveness, Fraud detection systems want to be adaptive. An associated situation typically classified below feasible predictiveness is the time required to locate fraudulent transactions. Certain structures require close to real-time indicators of suspicious transactions.

Prior research has already been done on a few factors of smart financial fraud detection. Initial fraud detection research targeted closely statistical models including logistic regression, in addition to neural networks [16], [17]. The researchers located that neural networks were used for financial programs including forecasting because 1988 [18]. In 1995, the primary anticipated financial declaration fraud was the usage of a back-propagation neural community [19]. In this paper, they compared techniques throughout a quantitative spectrum such as statistical and computational techniques including regression and neural networks [20]. In 1998, researchers used a neural community primarily based totally on different financial ratios and variables and discovered it compared favourably to discriminant evaluation and logistic regression [21]. In 2001 and 2002, they have done a few trendy evaluations on fraud detection, focusing mainly on statistical learning [22], [23], and investigated financial declaration fraud in depth [24]. Recent fraud detection studies have been some distance extra numerous in strategies studied, even though the previous techniques are nevertheless popular. In 2004, they reviewed the look at trendy fraud detection through the usage of analytic techniques such as neural networks [25]. In this paper, they investigated a unique technique the usage of the game principle in 2005, which modelled fraudsters and detection techniques as opposing gamers in a sport, every striving to achieve the best financial advantage [26]. They studied healthcare fraud through the usage of a system mining technique [27].

In 2007, they studied logistic regression with coverage fraud, targeting a database of Spanish car coverage claims [28], [29]. Researchers as compared statistical strategies with neural networks to pick out fraudulent Greek production organizations [6] and targeted class and regression trees to remedy

financial declaration fraud in a choice of Chinese organizations [30]. Also, in 2007 delivered a genetic set of rules on Accounting and Auditing Enforcement Releases to come across fraudulent organizations in the US [17] and evaluate present fraud detection literature. They claimed that the most effective hit strategies of fraud detection to date, in addition to the maximum generally researched, have been class-primarily based totally [16]. Researchers used decision trees to look at financial declaration fraud for a choice of Chinese groups in 2008 [31]. They took a statistical technique to cover fraud detection, the usage of the equal samples that have been used previously [32]. Both researchers checked out visualizing credit score card fraud with self-organizing maps, that specialize in real-global samples from the Singaporean department of a global bank [8]. They changed the usual synthetic immune system technique with a coevolutionary technique, the usage of it to remedy transactional fraud with the automated teller and point-of-sale information for a financial organization [33].

In 2009, applied a combination of text mining and Bayesian perception networks to pick out disgruntled personnel probable to dedicate company fraud [34]. This paper mixed a Dempster-Schaefer adder with a Bayesian learner to remedy credit score card fraud with their very own synthesized information [11]. Sánchez et al. targeted credit score playing cards supplied with the aid of using a multinational branch store, and the usage of self-organizing maps to cluster and visualize fraudulent patterns [10]. In this newsletter, they as compared help vector machines with decision trees in fixing credit score card fraud, with a focal point on aggregating not unusual place transactional variables to create new inputs [35]. In 2010, studied Accounting and Auditing Enforcement Releases (AAER) with their very own textual content mining and help vector device hybrid to are expecting economic declaration fraud in US groups [36].

In 2011, as compared the capacity of logistic regression, help vector machines, and random forests on a massive pattern of credit score card transactions to pick out which have been fraudulent [7]. Both researchers mixed the strengths of genetic algorithms and scatter seek to create their very own hybrid technique. They used it to tune customer



spending with a massive Turkish bank, as a resource to predict the incidence of credit score card fraud [11]. In this paper, they created text-mining hybrids with the aid of using making use of different not unusual place strategies to behave because of the classifier. With a help vector machine, decision tree, and Bayesian belief network they controlled to effectively perceive fraud in the company's 10-K report filings [37]. Both researchers additionally studied sections of 10-K files for US groups recognized to be fraudulent, processing the text with a novel validation decomposition vector to categorize the samples [38]. They carried out system mining to the inner logs created with the aid of using a European financial institution to come across company fraud [39] and did a huge evaluation of present fraud detection [7]. Also, in 2011, as compared a massive variety of techniques to discover financial declaration fraud inside Chinese organizations. In addition, to helping vector machines, they checked out genetic programming, logistic regression, organization technique of information handling, and a lot of neural networks [13]. This newsletter created a universal framework for financial declaration fraud detection through the usage of response floor methodology [4], then in 2012 with the aid of using making use of an artificial immune system to expect credit score card fraud for a first-rate Australian bank [40].

In 2013 Huang investigated financial declaration fraud in a chain of Taiwanese organizations through the usage of logistic regression and a help vector machine [41]. Both scientists took an extra direct technique and targeted the litigation phase of the Securities and Exchange Commission website, making use of their very own text-mining set of rules to categorize financial declaration fraud [42]. In this paper, they studied the capacity of decision trees to pick out fraudulent credit score card transactions, and the usage of a six-month pattern from a first-rate bank [43]. In 2014 researchers used text mining to look at AAERs for Chinese groups that have been buying and selling publicly in the US [44]. researcher visualized credit scorecard fraud with self-organizing maps, focusing most effectively on accounts held with the aid of using citizens of Warsaw, Poland [45] researchers applied an artificial immune system to pick out credit scorecard

fraud for a nameless Brazilian bank [46] and investigated the prevailing kingdom of fraud detection studies [47].

In 2015, the data mining techniques are mentioned for fraud detection, which is based on the kind of consumer clustering and for every cluster representing a certain kind of consumer, the system could have distinct behaviour. Finally, also studied through a decision tree set of rules and a neural network model. Models can extract numerous policies associated with consumer behaviour which are chosen withinside the corresponding table and have a chance per cent to discover the suspected cases [48]. In 2016, k-means clustering is used for credit card fraud detection. Data is growing haphazardly for credit cards and the k-means set of rules is used for coming across transactions whether it is fraud transaction or a valid transaction [49]. In 2017, researchers checked various detection techniques primarily based totally on credit cards in phrases of Parameter Speed of detection and provides a survey of diverse techniques utilized in credit card fraud detection and prevention [50]. In 2018, there are two main focuses, first on fraud instances that cannot be detected primarily based totally on preceding records or supervised learning and secondly producing a model of deep Auto-encoder and restricted Boltzmann machine (RBM) that may reconstruct regular transactions to search out anomalies from regular patterns [51].

In 2019, this paper could be very vital for ATM card issuers to select the best optimum solution for fraud detection problem, additionally permit us to construct a hybrid technique for growing a few optimum algorithms that can carry out nicely on fraud detection mechanism [52]. In 2020, numerous techniques have been developed primarily based totally on Artificial intelligence, Machine learning, Data mining, Genetic programming, Fuzzy logic etc. for detecting credit card fraudulent activities. On the other hand, the K-Nearest Neighbour algorithm and outlier detection techniques are applied to optimize the exceptional answer for the fraud detection problem [53]. In 2021, develop a model to analyze the imbalanced credit card fraud dataset [54].

II. Fraud Detection and Prevention

Fraud Detection and Prevention is a system-installed software program that may analyze any

inappropriate activity, offering risk mitigation and safety monitoring. It differs from a network protection strategy and enables the computer to identify suspicious activity before theft or other crimes are committed. Tools for detecting and preventing fraud are used as investigative techniques to find and stop fraud on a company device. These algorithms examine data from many different sources to look for probable errors like anomalies or illusions. It is utilized by a variety of businesses and organizations, including those in the life sciences, healthcare, travel, and government work. It is used to prevent cybercrimes that harm a company or organization, including account theft,

malware, hacking, DDoS, phishing, and credit card identity theft.

A. Fraud Detection

The system of the fraud detection which can be detect and also manages scammers from getting cash or possessions means wrong. It is a collection of actions designed to disclose and prevent attempts from scammers to acquire money or assets fraudulently. Fraud detection is popular in banks, insurance, healthcare, government, and the public sectors, in addition to regulation enforcement agencies. The overall working scenario is shown in Figure 2.

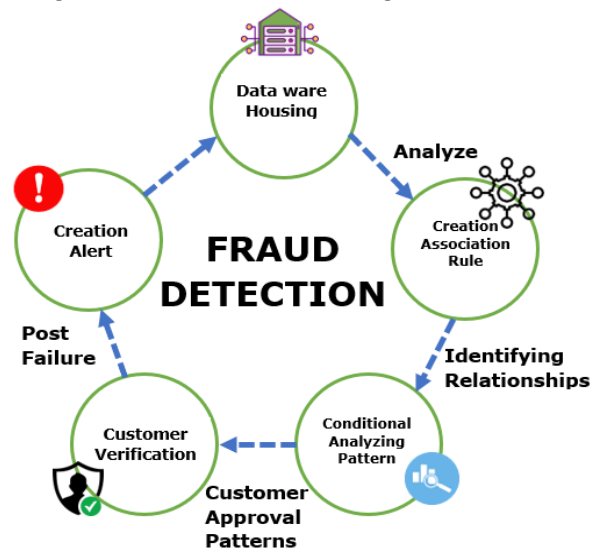


Figure 2: Fraud Detection Working Scenario

Laundering of money, cyberattacks, fake monetary privileges, bogus financial-bank cheques, burglary identification and lots of greater illegal moves that are the cases of duplicitous movement. From now, to respond the upward thrust in deceitful dealings throughout diverse stages, corporations adapt front-line fraud detection and prevention methods in addition to the strategies of risk management.

B. Types of Fraud Detection Techniques

The techniques based on data analysis are generally applied to detect fraud. The particular approaches

may be roughly grouped into different categories such as artificial intelligence-based, and statistical data analysis or computational methods. Imagine artificial intelligence, machine learning, neural networks, and deep learning as russian-nesting dolls [55] as shown in Figure 3. This is maybe the simplest way to conceptualize these concepts. Every one of them functions as a part of the previous work.

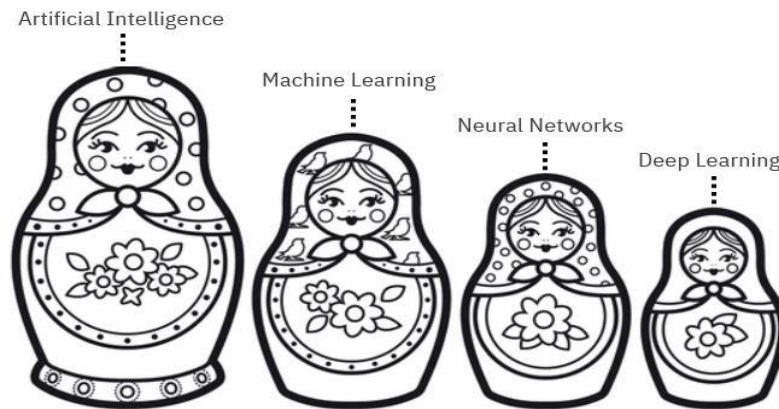


Figure 3: Count Nested

In other words, artificial intelligence includes the field of machine learning. The algorithms of deep learning which is the foundation of the neural networks and these are the branch of machine learning. In actuality, the neural network depth, the node layers having wide variety that splits it from a deep learning approach are requires greater than three layers. The detection techniques as given below,

a) **Artificial Immune System (AIS)**

The data mining strategy is the artificial immune systems which detects antigens through mimicking the biological immune system behavior [56]. The artificial immune system may imitate a wide range of biological traits, but the majority of them revolve around the detector cells formation and having potential capability to recognize external things. The cells of detector are created at random, and reproduction is used to check and assess their efficacy, in comparable with how other classification systems train.

Clonal selection is a typical kind that produces the cells of detector which at most exist for the brief

period. When a cell identifies an antibody that lives longer to combat the invader and can transform such as an outcome of the battle. The cells that survive after the imitation are the finest prepared to recognize the antitoxins. Negative selection is one more frequent method that the whole thing at random producing cells and defining how they interact with another epidemic cells in the system. In general, it has deleted and leaving the remaining capable of detecting intruders [57].

b) **Neural Network (NN)**

This is a computer model of the human brain which is named as neural network that represents neuronal and synapsis using the vertices and edges graphs [3]. The network works by modeling the enter variables as a layer of vertices after which making use of a weight to every link withinside the graph, whereas the ultimate vertices are located at distinctive layers primarily based totally on their distance from the enter nodes [58] as shown in Figure 4.

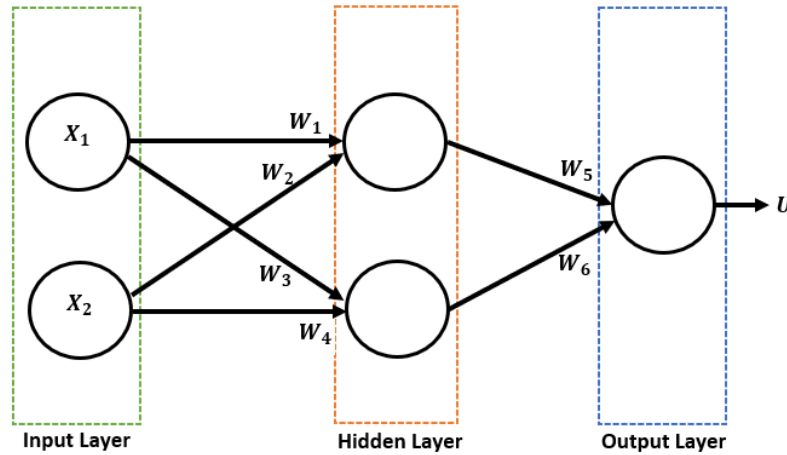


Figure 4: Simple Neural Network

Each node bases its input on the associated vertices to it preceding layer. The received signal through each neuron j is given by

$$U_{ij} = \sum W_{ij} \times X_i$$

Where W_{ij} shows the link weight of neurons i and j and X_i represents the input. If the result exceeds a certain threshold, the existing neuron fires and develops an input for the following layer.

Training a back proliferation neural network involves putting trials from the training data set through the system and compared the outcomes. At each edge weights are generally selected arbitrarily for first iteration, and when the results are computed, every weight is lightly changed through the following sequence [59]. The process is repeated until either the network's error has been decreased to an acceptable level or a predefined iteration limit has remained achieved. Following iteration, the network's performance can be evaluated using a set of validation data [3]. Overtraining is a typical issue through backpropagation the neural network, causing the network to emphasis on trends specific to the set of training data rather than broader challenge [59].

c) Genetic Algorithm (GA)

To iteratively enhance issue solutions, genetic algorithms employ the notion of resident development. It works by establishing a beginning group at random, then repeatedly replicating every resident utilizing various methods and choosing survivors depending upon their strength. Reproduction is accomplished by taking two exiting generation parents and employing crossover on dual places, at that time arbitrarily transforming a individual element of the resultant successors. A fitness function is used to assess the capacity of the offspring, and the results determine whose parents and kids are chosen as the future generation's representatives. The proportion of samples that the kids properly classify can be utilized to gauge their level of strength. The method finishes when it achieves the desired strength, even though to prevent indefinite looping, a limit on the number of iterations can be stated, as illustrated in Figure 5. Similar to neural networks, genetic algorithms may uncover underlying correlations among the data without the need for advance information of the issue domain. [60].

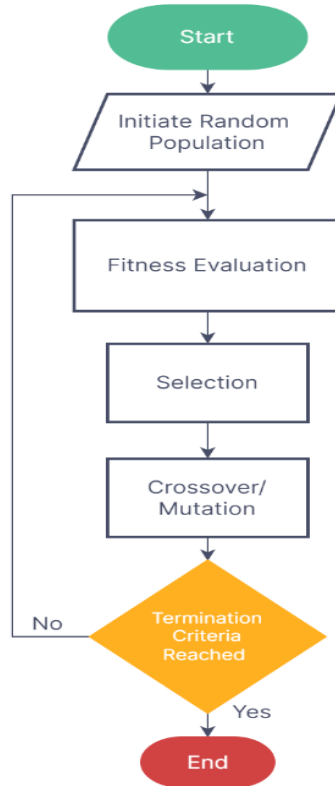


Figure 5: Flowchart of Genetic Algorithm Process

d) **Hidden Markov Model (HMM)**

The is a statistical model named as hidden markov model in which the represented system is believed to remain a markov process through an unseen state [61]. It detects fraud by analyzing user spending profiles, which are classified into three types [62]:

- I.lower profile
- II.middle profile
- III.higher profile

Figures 6 depict the training and detection and preventive phases [63] [64] of the procedure. In this

setup, launch the bank server and the HMM server first. When a transaction is initiated by the client, HMM begins watching and comparing the process. If fraud is detected, the transaction is stopped. The user responds with a password on a cellphone through Bluetooth to the similar ATM of bank, otherwise via message/sms. The passkey is validated for authorization, and the transaction is permitted. After three failed tries, the transaction is completely halted.

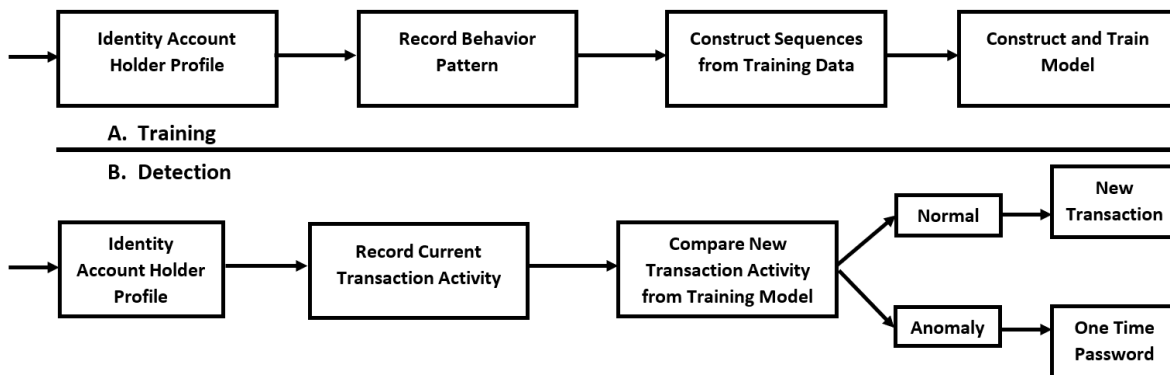


Figure 6: Flow chart of Training and Detection Phase in HMM

e) **Bayesian Belief Network (BBN)**

A statistical categorization approach is the Bayesian belief network which employs the theorem of Bayes, which is a way of determining the likelihood having a given hypothesis is true. According to the theorem, the probability P for a hypothesis H . For example, χ may be categorized inside a certain type that is specified as

$$P(\chi|H) = \frac{P(\chi|H)P(H)}{P(\chi)}$$

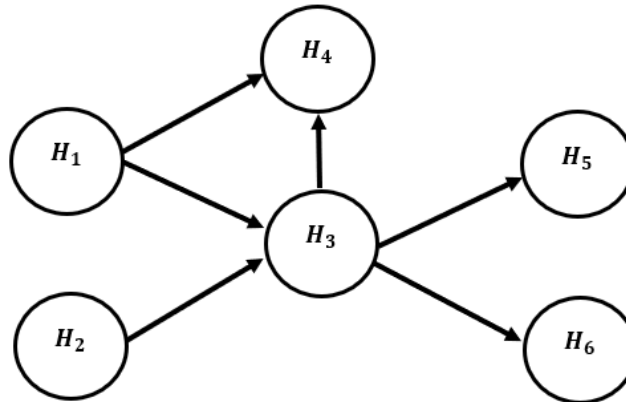


Figure 7: BBN Graphical Representation

A network uses a classifier to calculate $P(C_i|\chi)$ for all possible classes C_i and inserts χ into the class with the highest $P(C_i|\chi)$. In this way, the network is demonstrated to categorize each sample into the class to which it is most likely to belong [6].

A network may be represented graphically as a focused on acyclic graph, having nodes which represent trials and edges representing a fundamental relationship among them as shown in Figure 7. The absence of missing edges may thus be used to exhibit in which two variables are unrelated of each other [3].

f) **Cluster Method (CM)**

The cluster method is the procedure of organizing information within classes of things that are alike. Several cluster algorithms existing in classes of the dataset produce different grouping outcomes. The

approach used will be determined by the intended outcome [65]. The clustering of k-means is a modest and effective approach to data clustering. Figure 8 depicts the clustering-based approach's system architecture [66].

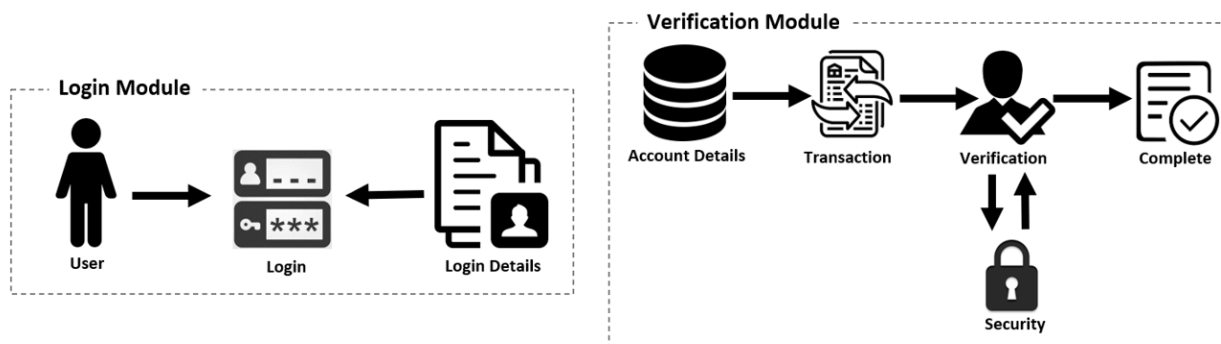


Figure 8: System Architecture of Cluster Method

Firstly, the parameters utilized in the programme, just like transaction award, credit/debit card number, current transaction, transaction time, mercantile group id, transaction category id and transaction state, are declared. The validation

mechanism then verifies the accuracy of the transaction information. The previously prepared data table is now inserted within the database. The information that is being removed from counter and now inserted to take transaction info. The

transaction information is then produced row by row using an array. Following that, the cluster is labelled as down, up, or moderate dangerous. The current transaction information was obtained to detect fraud or real transactions using the k-means clustering method. Uncertainty, the transaction is deceitful, the notification says "fraud transaction," or else it will say "legal transaction."

g) Self-Organizing Map (SOM)

The main type of artificial neural network is self-organizing map that consists of a single neural matrix. Inputs from a high-dimensional space are mapped to a two-dimensional array of neurons, a non-linear method is utilized. The mapping is intended to model comparable input vectors as neurons which are nearer together in the final matrix, allowing the inputs to be seen. To group the nodes, a distance or neighborhood function, just

like the euclidean distance formula or the gaussian formula, is utilized [63]. The clustering function that each neuron is subjected to is provided by:

$$y_{i+1} = y_i + \alpha(x_i - y_{i-1})$$

Where y_i represents the specific node present weighting, x_i represents the present input vector and α represents the preferred the function of distance. Before the process is finished, the clustering phase is repeated a predetermined number of times [67].

h) Decision Tree (DT)

Decision trees are a classification or prediction approach that utilizes a tree with internal nodes reflecting binary options on characteristics and branches expressing the result of that decision [59] as shown in Figure 9. Now trial travels the tree, it is segregated within the subsets though it is finally categorized within a jointly special subclass.

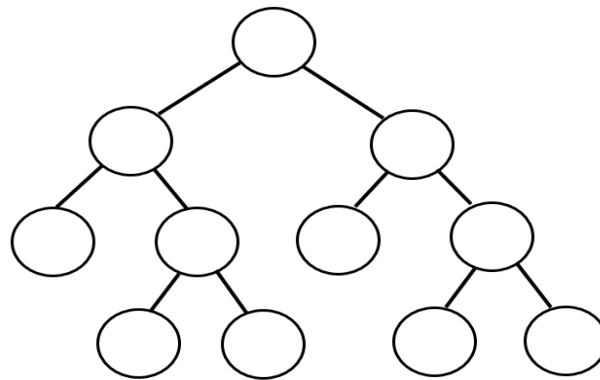


Figure 9: Decision Tree Representation

A decision forest, sometimes known as a random forests that is a decision tree collection intended to prevent the unstableness and risk of exaggerate in which an individual tree may cause [5]. Random forests employ distinct training info among tree and limit the characteristics pool presented to each internal node at random [5]. Pruning is another strategy for decreasing overfitting in decision trees, which includes removing decision nodes without affecting the tree's overall accuracy [6]. These approaches render random forest resistant to exaggerate and noise. Because every tree is created randomly, the computing complexity is minimal. Furthermore, the only two factors that must be adjusted are the number of trees and the collection of characteristics from which to create each node,

making decision forests straightforward to generate [5].

i) Super Vector Machine (SVM)

Support Vector Machines (SVM) are statistical learning approaches that have been successfully used for a variety of issues. The essential idea behind the SVM sorting method is to build a hyperplane known the decision plane, maximizing the distance among the positive and negative modes [68]. SVM is a well-known machine learning approach for sorting, regression, and additional problems. LIBSVM is a Support Vector Machines library (SVM). LIBSVM is often used in two stages: first, setting a training sequence to generate a model. After that utilizing the model to guess information from a



testing info set. The SVM have major functions are as follows:

- First, set up the training information for model development.
- Next, arrange SVM elements for the newly produced dataset and refer it to SVM training.
- SVM Trainer, which trains every single data point in the big dataset.
- After the dataset has been entirely trained, the SVM Predictor predicts the learned data.

C. Supervised and Unsupervised Learning

There are important strategies utilized in machine learning and artificial intelligence which are supervised and unsupervised learning. One employ labelled data to support in outcome estimation, at the same time as the opposite does not. This is a significant difference and there are several factors areas where one of the two approaches performs better than the other, although the two strategies do differ somewhat from one another.

a. Supervised Learning

The supervised learning method to machine learning is prominent by the utilize of labelled datasets. Classification and regression are two main types that may be used to classify supervised learning when applying data mining.

i) Classification

The Classification issues utilize an algorithm to exactly distribute test data into various classes, just like distinctive among apples and oranges. An alternative is to segregate spam from your email in a separate folder using supervised learning techniques. Decision trees, support vector machines, random forests, and linear classifiers are examples of common classification approaches.

ii) Regression

An algorithm is utilized in regression, a distinct supervised learning technique, to get the relationship among dependent and independent parameters. Regression models are beneficial while expecting numbers based on a species of data sources, just like sales revenue estimations for a particular organization. Logistic, linear and polynomial regressions are a few mutual regression methods.

b. Unsupervised Learning

Unsupervised learning investigates and classifies unlabeled statistical sets using machine learning methods. Without the support of humans, these algorithms look for statistics that point to hidden elegances. Clustering, association, and dimensionality lessening were the three main functions applied in unsupervised learning paradigms.

i) Clustering

Unlabeled data may be characterized utilizing the statistics mining method, in which clusters objects only based upon their resemblances or alterations. K-means clustering approaches split associated statistical elements within clusters based upon comparisons, and the k-means identifies the dimensions and amount of granularity of the clusters. Various parameters, consisting market segmentation and image reduction, build this tactic attractive.

ii) Association

The association shape of unsupervised learning utilizes a diversity of measures to novelty associations among parameters in a provided dataset. Both the marketplace basket evaluation and the "Customers who bought this item also bought" recommendation engine frequently utilize those approaches.

iii) Dimensionality Reduction

Dimensionality reduction is a learning method utilized when a dataset has excessive properties (or dimensions). It mitigates the info inputs volume to a well-behaved level whereas retaining the data integrity. This method is broadly utilized for pre-processing data, just like whenever autoencoders enhance the video quality of resultant images.

D. Fraud Prevention

Corporations and politicians have adopted technologies like data analytics and artificial intelligence to significantly reduce and even avoid the economic, social, and financial consequences of fraud. Consequently, analysts and researchers eliminate barriers, discover and rank severity-based alerts, and then present high-priority indications for further analysis.

Advances in fraud detection technology serve as a precise and effective weapon against scammers.

There are eight steps for fraud prevention as shown in Figure 10.



Figure 10: Fraud Prevention Steps

The fraud prevention strategy presents a high-level plan for executing the Institution's fraud prevention policy. In view of the fact that the approach is the supreme important factor of the fraud prevention plan, it should be simple and realistic. The fraud risk management policy and the fraud risk profile of the institution determine the fraud prevention approach. There are some fraud prevention strategies as follows [69] [70].

a) Identification and Evaluation of Sensitive Regions

To develop and implement a fraud prevention strategy, the organization should firstly define wherever fraud risks exist in the Institution's present operational systems and processes. Only when these exposures have been recognized will it be feasible to take corrective action and, if possible, avoid or minimize the occurrence of fraud in the future.

b) Fraud Risk Ownership

To some extent, all staff are accountable for managing fraud risk, although the Accounting Officer / Authority has final accountability. Line managers in certain areas of the Institution may be delegated authority by the Accounting Officer / Authority. The Accounting Officer / Authority has the authority to transfer responsibilities for fraud risk management as well as the flow of operations from the strategic to the operational level.

c) Plan of Action

The Institution should define clear processes for dealing with control deficiencies as part of the response plan. The organization must establish

clean reporting forms for scam. Scam reporting must be incorporated into the respond strategy or investigative policy. The rejoinder proposal must include define the actions and persons in charge of each response activity.

d) The Legal Framework

The necessary legislation for dealing with civil and criminal offences against the Institution should be defined and properly construed. It should be obvious what defines a fraudulent or corrupt behaviour.

e) Culture of Anti-Fraud

The Accounting Officer / Authority should establish frameworks to promote and educate stakeholders about the Institution's anti-fraud and anti-corruption culture. As part of an anti-fraud strategy, management might be entrusted with the obligation of teaching other employees under their supervision on fraud and corruption.

III. Analysis and Discussion

We will categorize the financial fraud detection techniques described in this area based on their success rate, the method used, and the fraud type analyzed. This classification will allow us to illustrate patterns in existing research methodologies. The study's objectives were to ascertain the operational response of various fraud detection methods. We conducted a comparative study on fraud detection methods to analyze the results. For comparison, we considered the most important parameters such as accuracy, speed, and cost. A comparison table has been created to

compare different ATM card fraud detection mechanisms. Each of the ATM card fraud detection techniques described in this study has its own set of benefits and drawbacks. Table 1 shows the comparison results obtained from this study [71], [72], [73], [74],[75], [76].

This study examined the performance of various ATM cards fraud detection techniques such as neural networks, genetic algorithms, Hidden

Markov models, Bayesian networks, decision trees, clustering methods, support vector machines (SVM), and artificial immune systems. As a result, each method has benefits and drawbacks. At the same time, the support vector machine has a low detection speed and the artificial immune system has a high detection speed. So, based on the results, the best method among these techniques is AIS, NN, GA, HMM, BBN, CN, SOM, DT and SVM.

Table 1: Benefits and Drawbacks of various Fraud Detection Techniques

Techniques	Benefits	Drawbacks
Artificial Immune System (AIS)	Self-organization, ease of integration with other systems, and fault tolerance	In the NSA, extensive training is required.
Neural Networks (NN)	High detection accuracy, portability, and speed	High cost/data format sensitivity
Genetic Algorithm (GA)	Detection is inexpensive and quick.	Setup and operation require extensive tool knowledge and are difficult to understand.
Hidden Markov Model (HMM)	Rapid detection.	Low accuracy/incapability to handle large data sets
Bayesian Belief Network (BBN)	To operate, data must be trained and a high processing speed is required. More precise and faster than a neural network.	Excessive training is required, and BBN's are slower to apply to new instances.
Clustering method (CN)	Clustering assists in grouping data into similar clusters, allowing for simple data retrieval.	Numerous non-fraudulent activities were mistakenly identified as frauds. So, to detect fraud accurately and efficiently, real data must be available.
Self-Organizing Map (SOM)	To reduce incredibly complex problems down to easily interpreted data mappings	Requires neuron weights to be necessary and sufficient to cluster inputs.
Decision Tree (DT)	High adaptability/ease of implementation	Each condition must be checked individually. The transaction condition is used in fraud detection.
Support Vector Machine (SVM)	SVMs can be robust even if the training sample is biased.	Expensive/poor performance when processing large datasets.

This paper presents a comparative study of various credit card fraud detection techniques. The primary

goal of this paper is to examine the methodology of various credit card detection methods. The survey-



based comparison of the mentioned approaches in terms of parameters such as detection speed, accuracy, and the cost is shown in Table 2.

Table 2: Comparison of various Fraud Detection Techniques

Techniques	Detection Speed	Accuracy	Cost
AIS	Very Fast	Good	Inexpensive
NN	Fast	Medium	Expensive
GA	Good	Medium	Inexpensive
HMM	Fast	Low	High Expensive
BBN	Very Fast	High	Expensive
CM	High	Medium	Expensive
SOM	Fast	Medium	Expensive
DT	Fast	Medium	Expensive
SVM	Low	Medium	Expensive

IV. Conclusion

Fraud detection is an essential component of the modern financial business. This study of the literature focused on research on statistical and computational intelligence techniques for fraud detection. Despite differences in effectiveness, each approach was demonstrated to be relatively capable of identifying various types of financial fraud. The capacity of computational approaches such as neural networks and support vector machines to learn and adapt to new strategies is extremely useful in fraudsters growing strategies.

The main objective of this work is to review various fraud detection methods. Fraud detection and prevention should be a key concern for every company. A well-planned and implemented fraud detection system may lower the likelihood of fraud occurring inside a company dramatically. Furthermore, the quick discovery of fraud has a direct beneficial impact on the firm by lowering future potential losses. AI and statistical data analysis are effective detection approaches that act as a deterrent to potential scammers. As legal requirements and regulatory demands have increased, it has become more critical to create an effective fraud detection and prevention program. All the fraud detection techniques presented in this project have both strengths and disadvantages. Some approaches have a high detection speed but a low accuracy. Some approaches offer high accuracy but are prohibitively costly.

V. Future Work and Challenges

Although data-driven artificial intelligence systems have demonstrated remarkable performance in the detection of financial fraud, significant concerns remain unresolved as financial fraud schemes evolve to adapt to this new digital environment. As follows, we present the primary problems and offer future work directions from task-oriented, data-oriented, and model-oriented perspectives.

- Financial fraud is becoming more difficult to detect due to its increasing secrecy and complexity.
- The secrecy of financial fraud causes natural inaccuracy in sampling.
- The intricacy of financial processes necessitates the involvement of large amounts of data.
- The amount of financial data available for fraud detection is vast, yet it is dispersed.
- Data isolation is a challenging problem to address.
- Model training is made more difficult by large-scale data processing.
- Models for detecting financial fraud must be more adaptable and interpretable.
- The issue of model bias must be addressed.
- Robustness should be improved.
- Improved interpretability is required.



REFERENCES

- <https://legaljobs.io/blog/credit-card-fraud-statistics/> - Jenifer Kuadli - 2022
- <https://www.fool.com/the-ascent/research/identity-theft-credit-card-fraud-statistics/> - Lyle Daly and Jack Caporal, 2022.
- Ngai E., Hu Y., Wong Y., Chen Y., and Sun X., "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," *Decision Support Systems* 50, 559-69, 2011.
- Zhou W. and Kapoor G., "Detecting evolutionary financial statement fraud," *Decision Support Systems* 50, 570-5, 2011.
- Bhattacharyya S., Jha S., Tharakunnel K., and Westland J. C., "Data mining for credit card fraud: A comparative study," *Decision Support Systems* 50, 602-13, 2011.
- Kirkos E., Spathis C. and Manolopoulos Y., "Data mining techniques for the detection of fraudulent financial statements," *Expert Systems with Applications* 32, 995-1003, 2007.
- Ngai E., Hu Y., Wong Y., Chen Y. and Sun X., "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," *Decision Support Systems* 50, 559-69, 2011.
- Quah J. T. and Sriganesh M., "Real-time credit card fraud detection using computational intelligence," *Expert Systems with Applications* 35, 1721-32, 2008.
- Yeh I. and Lien C. H. "The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients," *Expert Systems with Applications* 36, 2473-80, 2009.
- Sánchez D., Vila M., Cerda L. and Serrano J. M., "Association rules applied to credit card fraud detection," *Expert Systems with Applications* 36, 3630-40, 2009.
- Panigrahi S, Kundu A, Sural S, and Majumdar AK, "Credit card fraud detection: A fusion approach using Dempster Shafer theory and Bayesian learning," *Information Fusion* 10, 354-63, 2009.
- Duman E. and Ozelik M. H., "Detecting credit card fraud by genetic algorithm and scatter search," *Expert Systems with Applications* 38, 13057-63, 2011.
- Ravisankar P., Ravi V., Raghava Rao G., and Bose I., "Detection of financial statement fraud and feature selection using data mining techniques," *Decision Support Systems* 50, 491-500, 2011.
- Judith Hurwitz, Alan Nugent, Fern Halper, and Marcia Kaufman, "How Big Data Analytics Can Prevent Fraud," chapter 22, page 260. *Big Data for Dummies*. John Wiley & Sons, 2013.
- Richard J. Bolton and David J. Hand, "Statistical Fraud Detection: A Review," *Journal of Statistical Science*, 17:235-255, 2002.
- Yue D., Wu. X., Wang Y., Li Y., and Chu C. H, "A review of data mining based financial fraud detection research," In *Wireless Communications, Networking and Mobile Computing, WiCom. International Conference on*. (ed.), Vol. pp. 5519-22, IEEE, 2007.
- Hoogs B., Kiehl T, Lacombe C., and Senturk D., "A genetic algorithm approach to detecting temporal patterns indicative of financial statement fraud," *Intelligent Systems in Accounting, Finance and Management* 15, 41-56, 2007.
- Zhang G., Eddy Patuwo B., and Y. Hu M., "Forecasting with artificial neural networks: The state of the art," *International journal of forecasting* 14, 35-62, 1998.
- Sohl J. E. and Venkatachalam A., "A neural network approach to forecasting model selection," *Information & Management* 29, 297-303, 1995.
- Fraser I. A., Hatherly D. J., and Lin K. Z., "AN EMPIRICAL INVESTIGATION OF THE USE OF ANALYTICAL REVIEW BY EXTERNAL AUDITORS," *The British Accounting Review* 29, 35-47, 1997.
- Fanning K. M. and Cogger K. O., "Neural network detection of management fraud using published financial data," *International Journal of Intelligent Systems in Accounting, Finance & Management* 7, 21-41, 1998.



- Bolton R. J. and Hand D. J., "Statistical fraud detection: A review," *Statistical Science* 235-49, 2002.
- Bolton R. J. and Hand D. J., "Unsupervised profiling methods for fraud detection," *Credit Scoring and Credit Control VII* 235-55, 2001.
- Rezaee Z., "In *Financial statement fraud: prevention and detection*," Vol. pp. John Wiley & Sons, 2002.
- Kou Y., Lu C. T., Sirwongwattana S., and Huang Y. P., "Survey of fraud detection techniques," In *Networking, sensing and control, IEEE international conference on*. (ed.), Vol. 2, pp. 749-54, IEEE, 2004.
- Vatsa V., Sural S., and Majumdar AK., "A game theoretic approach to credit card fraud detection," In *Information Systems Security*. Vol. pp. 263-76. Springer, 2005.
- Yang W. S. and Hwang S. Y., "A process mining framework for the detection of healthcare fraud and abuse," *Expert Systems with Applications* 31, 56-68, 2006.
- Pinquet J., Ayuso M., and Guillen M., "Selection bias and auditing policies for insurance claims," *Journal of Risk and Insurance* 74, 425-40, 2007.
- Viaene S., Ayuso M., Guillen M., Van Gheel D., and Dedene G., "Strategies for detecting fraudulent claims in the automobile insurance industry," *European Journal of Operational Research* 176, 565-83, 2007.
- Bose I. and Wang J., "Data mining for detection of financial statement fraud in Chinese Companies," *International Conference on Electronic Commerce, Administration, Society and Education*, Hong Kong, 2007.
- Bai B., Yen J., and Yang X., "False financial statements: characteristics of China's listed companies and CART detecting approach," *International Journal of Information Technology & Decision Making* 7, 339-59, 2008.
- Bermúdez L., Pérez J., Ayuso M., Gómez E., and Vázquez F., "A Bayesian dichotomous model with asymmetric link for fraud in insurance," *Insurance: Mathematics and Economics* 42, 779-86, 2008.
- Wu S. X. and Banzhaf W., "Combatting financial fraud: a coevolutionary anomaly detection approach," In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*. (ed.), Vol. pp. 1673-80, ACM, 2008.
- Holton C., "Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion-dollar problem," *Decision Support Systems* 46, 853-64, 2009.
- Whitrow C., Hand D. J., Juszczak P., Weston D., and Adams N. M., "Transaction aggregation as a strategy for credit card fraud detection," *Data Mining and Knowledge Discovery* 18, 30-55, 2009.
- Cecchini M., Aytug H., Koehler G. J., and Pathak P., "Making words work: Using financial text as a predictor of financial events," *Decision Support Systems* 50, 164-75, 2010.
- Humpherys S. L., Moffitt K. C., Burns M. B., Burgoon J. K., and Felix W. F., "Identification of fraudulent financial statements using linguistic credibility analysis," *Decision Support Systems* 50, 585-94, 2011.
- Glancy F. H. and Yadav S. B., "A computational model for financial reporting fraud detection," *Decision Support Systems* 50, 595-601, 2011.
- Jans M., vander Werf J. M., Lybaert N. and Vanhoof K., "A business process mining application for internal transaction fraud mitigation," *Expert Systems with Applications* 38, 13351-9, 2011.
- Wong N., Ray P., Stephens G. and Lewis L., "Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results," *Information Systems Journal* 22, 53-76, 2012.
- Huang S. Y., "Fraud Detection Model by Using Support Vector Machine Techniques," *JDCTA: International Journal of Digital Content Technology and its Applications* 7, 32-42, 2013.
- Zaki M. and Theodoulidis B., "Analyzing Financial Fraud Cases Using a Linguistics Based Text Mining Approach," *Available at SSRN* 2353834, 2013.



- Sahin Y., Bulkan S., and Duman E., "A cost-sensitive decision tree approach for fraud detection," *Expert Systems with Applications* 40, 5916-23, 2013.
- Dong W., Liao S. S., Fang B., Cheng X., Chen Z., and Fan W., "The Detection of Fraudulent Financial Statements: An Integrated Language Model," 2014.
- Olszewski D., "Fraud detection using self-organizing map visualizing the user profiles," *Knowledge Based Systems*, 2014.
- Soltani Halvaiee N. and Akbari M. K., "A novel model for credit card fraud detection using Artificial Immune Systems," *Applied Soft Computing*, 2014.
- West J., Bhattacharya M. and Islam R., "Intelligent Financial Fraud Detection Practices: An Investigation", 10th International Conference on Security and Privacy in Communication Networks, 2014.
- Singh, P. and Singh, M., "Fraud Detection by Monitoring Customer Behavior and Activities," *International Journal of Computer Applications*, 111, 23-32, 2015.
- Sonawane, Y.B., Gadgil, A.S., More, A.E. and Jathar, N.K., "Credit Card Fraud Detection Using Clustering Based Approach," *International Journal of Advance Research and Innovative Ideas in Education*, 2, 1773-1776, 2016.
- Gupta, Surbhi, Mrs. and Nitima Malsa. "Credit Card Fraud Detection & Prevention - A Survey," *International Journal for Innovative Research in Science & Technology*, vol.1, 4, 2017.
- Pumsirirat, A. and Liu, Y., "Credit Card Fraud Detection Using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine," *International Journal of Advanced Computer Science and Application*, 9, 18-25, 2018.
- Rahman, M. and Saha, A., "A Comparative Study and Performance Analysis of ATM Card Fraud Detection Techniques," *Journal of Information Security*, 10, 188-197, 2019.
- Pooja, Dr. Ashlesha, "Review on Credit Card Fraud Detection using Machine Learning Algorithms," *International Journal of Computer Trends and Technology*, 68.6, 77-81, 2020.
- Panda, A., Yadlapalli, B., & Zhou, Z., "Credit card fraud detection through machine learning algorithm," *Big data and computing visions*, 1 (3), 140-145, 2021.
- B. Jiang and Y. Mu, "Russian Doll Network: Learning Nested Networks for Sample-Adaptive Dynamic Inference," *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pp. 336-344, 2021.
- Wu SX and Banzhaf W, "Combatting financial fraud: a coevolutionary anomaly detection approach," In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*. (ed.), Vol. pp. 1673-80, ACM, 2008.
- Soltani Halvaiee N and Akbari MK, "A novel model for credit card fraud detection using Artificial Immune Systems," *Applied Soft Computing*, 2014.
- Koh HC and Low CK, "Going concern prediction using data mining techniques," *Managerial Auditing Journal* 19, 462-76, 2004.
- Zhang D and Zhou L, "Discovering golden nuggets: data mining in financial application," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 34, 513-22, 2004.
- Su, Jianhai & Havens, Timothy., "Fuzzy community detection in social networks using a genetic algorithm." *IEEE International Conference on Fuzzy Systems*. 2039-2046. 10.1109/FUZZ-IEEE.2014.6891611, 2014.
- Sonawane, V.D., Gupta, P., Raut, A. and Saudagar, F., "ATM Card Fraud Detection Using Hidden Markov Model," *International Journal of Innovative Research in Computer and Communication Engineering*, 4, 8742-8747, 2016.
- Patidar, R. and Sharma, L., "Credit Card Fraud Detection Using Neural Network," *International Journal of Soft Computing and Engineering*, 1, 32-38, 2011.
- Mhamane, S.S. and Lobo, L.M.R.J., "Use of Hidden Markov Model as Internet Banking Fraud Detection," *International Journal of Computer Applications*, 45, 5-10, 2012.



- Bhingarde, A., Bangar, A., Gupta, P. and Karambe, S., "Credit Card Fraud Detection Using Hidden Markov Model," *International Journal of Advanced Research in Computer and Communication Engineering*, 4, 169-170, 2015.
- Vaishali, "Fraud Detection in Credit Card by Clustering Approach," *International Journal of Computer Applications*, 98, 29-32, 2014.
- Sonawane, Y.B., Gadgil, A.S., More, A.E. and Jathar, N.K., "Credit Card Fraud Detection Using Clustering Based Approach," *International Journal of Advance Research and Innovative Ideas in Education*, 2, 1773-1776, 2016.
- Quah JT and Sriganesh M, "Real-time credit card fraud detection using computational intelligence," *Expert Systems with Applications* 35, 1721-32, 2008.
- Olszewski D, "Fraud detection using self-organizing map visualizing the user profiles," *Knowledge-Based Systems*, 2014.
- S. Surbhi and D. S. Kumar, "Fraud Detection During Money Transaction and Prevention," 2019 *International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, pp. 1-4, 2019.
- S. Hoyer, H. Zakhariya, T. Sandner and M. H. Breitner, "Fraud Prediction and the Human Factor: An Approach to Include Human Behavior in an Automated Fraud Audit," 2012 *45th Hawaii International Conference on System Sciences*, 2012, pp. 2382-2391, 2018.
- Zareapoor, M., Seeja, K.R. and Alam, M.A., "Analysis of Credit Card Fraud Detection Techniques: Based on Certain Design Criteria," *International Journal of Computer Applications*, 52, 35-42, 2012.
- Kumari, S. and Choubey, A. "A Review on Various Techniques and Approaches for Credit Card Fraud Detection," *International Journal of Scientific Research Engineering & Technology*, 6, 485-489, 2017.
- Bhatia, S., Bajaj, R. and Hazari, S., "Analysis of Credit Card Fraud Detection Techniques," *International Journal of Science and Research*, 5, 1302-1307, 2016.
- Singh, P. and Singh, M., "Fraud Detection by Monitoring Customer Behavior and Activities," *International Journal of Computer Applications*, 111, 23-32, 2015.
- Pumsirirat, A. and Liu, Y., "Credit Card Fraud Detection Using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine," *International Journal of Advanced Computer Science and Applications*, 9, 18-25, 2018.
- Gupta, S. and Malsa, N. "Credit Card Fraud Detection and Prevention—A Survey," *International Journal for Innovative Research in Science & Technology*, 4,1-7, 2017.