

PAPER • OPEN ACCESS

## Fraud Detection in Financial Statements using Text Mining Methods: A Review

To cite this article: Ajit Kr. Singh Yadav and Marpe Sora 2021 *IOP Conf. Ser.: Mater. Sci. Eng.* **1020** 012012

View the [article online](#) for updates and enhancements.

You may also like

- [An Efficient Techniques for Fraudulent detection in Credit Card Dataset: A Comprehensive study](#)  
Akanksha Bansal and Hitendra Garg
- [Feature engineering strategies based on a One-point Crossover for fraud detection on Big Data Analytics](#)  
M Soleh, E R Djuwitaningrum, M Ramli et al.
- [Credit card fraud detection using neural network and geolocation](#)  
Aman Gulati, Prakash Dubey, C MdFuzail et al.



245th ECS Meeting • May 26-30, 2024 • San Francisco, CA

Present your work at the leading electrochemistry & solid-state science conference.

Network with academic, government, and industry influencers!

Submit abstracts by December 1, 2023

[Learn more & submit!](#)



## Fraud Detection in Financial Statements using Text Mining Methods: A Review

*Ajit Kr. Singh Yadav<sup>#</sup>, Marpe Sora<sup>\*</sup>*

<sup>#</sup> Asst. Prof. Dept. of CSE, NERIST, Itanagar, Arunachal Pradesh &  
Research Scholar, Dept. of CSE, RGU, Itanagar, Arunachal Pradesh  
Email: ajityadav101@rediffmail.com

<sup>\*</sup> Asst. Prof. Dept. of CSE, RGU, Itanagar, Arunachal Pradesh  
Email: marpe.sora@rgu.ac.in

**Abstract:** In the financial industry, financial fraud is one of the ever-growing hazards with far concerns. Financial statements are the fundamental papers which replicate economic position of a corporation. Users of the financial information like public, creditors etc. are the major foundations of a decision-making process for financing stakeholders. Financial fraud has extremely damaged the sustainable growth of financial markets and enterprises. The amount of financial reporting fraud cases keeps on developing. Each incident is a thick hit to partners, banks, and financial specialists and it costs human progress significantly. One of the serious issues is to recognize the financial reporting fraud by utilizing formation of an active model. The aim of this paper is to identifying frauds using various text mining techniques and guard, the public's investments. This investigation will benefit auditors and financial governors.

**Keywords:** *Text Mining, Fraud Detection, Financial statements, Adaptive Crime, Computational Fraud detection model.*

### 1. Introduction

In the modern world, the text is one of the common means for exchanging data. Text mining is an interdisciplinary field which includes ordinary linguistic processing, information retrieval, web mining, computational linguistics, data extraction and data mining. From formless and semi-structured papers, automatically structured information were extracted in text mining. It has a very high commercial value. For evaluating huge assortments of formless documents, it is an advance innovation technique for the purposes of extracting and knowledge or non-trivial patterns [1]. Document files have several formats like text files or flat files and PDF files. These files are gathered from multiple sources such as message boards, newsgroups, emails, online chat, SMS, web pages and so on. Humans can easily handle the problems and have the ability to identify and apply linguistic patterns to text. But computers cannot easily handle such as spelling, contextual meaning, slang and variations [2]. Nevertheless, to know unstructured data, the language abilities permit us and computer capability to process the text at high speeds or in large volumes. For analysis task on unstructured data, text mining helps to the computer. Fraud Detection has great prominence to Financial Organizations. In worldwide, financial fraud is a critical commercial issue. Numerous kinds of fraud include in financial fraud such as health-care fraud, mortgage fraud, corporate fraud, safeties and commodities fraud, financial organisation fraud, and others. Fraud detection takes into action when preclusion of fraud fails. Fraud may be unaware so that fraud preclusion has failed. Therefore, detection of fraud must be used repeatedly [3].

Criminals often change their place when they get information about the usage of the fraud detection scheme in a certain process. Plenty of criminals don't aware of this kind of protection scheme which causes ease of fraud identification [4]. Application of early detection tools is very helpful but it is difficult to process to exchange ideas in fraud detection schemes. Public description of fraud detection scheme is not good. Because criminals will easily access all details so that it is important to make more



secure schemes [5]. Conventional detection methods are not good in terms of economy, security, time consumption, law, and computational complexity. Assurance claims, credit card businesses, tax reappearance claims and cell phones are some possible ways of crime. These kinds of criminalities make more loss to governments and businesses due to a liberal manner of security [6].

Fraud is additionally called an adaptive crime. Therefore, it needs some superior techniques of intelligent data investigation to prevent and detect it. In several areas of fraudulent crimes, they are delivering applicable and successful solutions. Due to data investigation, these methods belongs to data mining, text mining, machine learning, statistics and knowledge discovery in databases (KDD) [7]. Financial organizations and banks create huge volumes of detailed transaction data. Data mining includes some primary applications such as trend analysis, direct marketing, fraud detection and risk assessment of potential customers at banks. For forecasting, most of the requirements dominate in the financial field. With an extraordinary level of estimation, commodity prices and predicting of stock prices can great incomes [8]. In predicting, options and bond trading, portfolio management, the neural network techniques are used and also applied in acquisitions and mergers. A whole sympathetic of nature, consequence and fraudulent financial reporting activities needs a suitable explanation of financial statement fraud. Association of certified fraud examiners (ACFE) describes financial statement fraud as the intentional, misstatement or accounting information or omission of material facts. That is deceptive when considered with all the information made obtainable; it would cause the reader to modify her or his decision [9].

Financial statement fraud is known as when the managers of a company deliver false financial information. Financial statement audits guarantee that a company's financial reports are free from material misstatement and fraud. In this scheme, financial statement fraud is typically committed [10]. To fight fraudulent activities, organizations want to be prepared in a recent challenging economy. By a business administration, reporting fraud includes the modification of financial statement data to attain a fraudulent result [11]. By the top management of publicly traded companies, financial statement fraud is a deliberate and illegal act devoted. Three features of frauds are available such as rationalization or attitude, opportunity and pressure or motive. In every year, financial statement fraud costs millions of dollars to the world economy because of the failure of many companies [12]. It is a type of the management fraud which operates the financial information. In many organizations, prevention of financial statement fraud has one of the major concerns [13]. Financial reporting fraud contains activities such as reporting sales which did not happen. In the recent year, reporting income into the current year essentially fits. In the current year, reporting an expense or capitalizing expenses in the next year that should be reported [14].

For discovery and identification of fraud, fraudulent observations are generally applied. The observations are the result of error or circumvent procedures in the prevention of fraud. A measure to break fraud in the first place from happening is known as fraud prevention [15]. In an extra structured spreadsheet layout, feature selection and feature generation events are accomplished to characterize the amorphous text documents based on these recovered and pre-processed documents. Feature selection procedures need an extensive exploration of all subdivisions of features to recognize the significant features [16]. For the applications of the knowledge discovery process, Information Extraction, Information Retrieval, the Text Mining methods are integrated after the suitable collection of features. Mostly classification is used to classify each item from a set of data into one of the predefined set of classes [17]. Classification procedure contracts with the software that can acquire to organize the data substances into groups. This scheme is also known as a process of identifying a motivating, inferred and innovative data in KDD.

In developing countries, financial statement fraud is more severe and it is a serious economic and social problem [18]. In yearly and quarterly, a company needs stock exchange to distribute its financial statements such as revenue, statements of recollected earnings, profit and loss account and cash flow statements. By analysing, financial report of a company replicates its genuine financial health. In the company, shareholders can form a wise decision about financing [19]. Financial fraud accomplishments of money laundering and credit card fraud gradually increase in recent years. It causes the loss of individual and/or enterprises' assets. Due to the profit from fraud, they threaten the safety of a country which may go to terrorism. Financial fraud detection and sketching fraud is critical and needed. Fraud detection is a difficult challenge for the reason of transactions and intricate trading systems. Many fraud detection models work with attribute-value data points which are obtained from transactions data [20]. Majority of fraud detection studies are applying data mining procedures which focussed on organised information using numerical approaches.

## 2. A BRIEF REVIEW OF FRAUD DETECTION IN FINANCIAL STATEMENTS

Financial fraud detection was talked about by Huang *et al* [21] with inconsistency feature detection. This article gives the CoDetect innovation to play out the fraud detection on highlight and chart based likeness matrix promptly. The innovation exhibits another strategy to uncover the idea of commercial exercises from fraud examples to disbelieving assets. In addition, to recognize the fraud on a sparse matrix, the structure gives an increasingly interpretable way. Exploratory results are tried on manufactured and true datasets which demonstrates the recommended structure can successfully identify guarded features and the fraud decorations.

Hybrid data mining plan based financial statement fraud detection was portrayed by Yao *et al* [22]. The primary motivation behind this investigation is to enhanced financial fraud detection conspire to consolidate feature selection and machine learning classification. Re-enactment considers obviously demonstrated that irregular forest gives better execution thought about than different frameworks. XGBOOST gives the better execution in the two feature selection strategies. In light of the exploration work, two or five factors are progressively adequate in this investigation.

Utilizing machine learning, credit card fraud detection was proposed by Yee *et al* [23] in the data mining process. Based on anomalies and standardized data, data mining strategies were utilized to examine the patterns and features of suspicious and non-suspicious exchanges. Something else, the machine learning techniques were utilized to consequently anticipate the non-suspicious and suspicious transactions by utilizing classifiers. By learning the patterns of the data, the blend of data mining and machine learning approaches could recognize the genuine and non-genuine exchanges.

ATM card fraud detection framework was presented by Prakash *et al* [24] utilizing machine learning strategies. This plan lessening the measure of false alert and can demonstrates accurate in deducting fraudulent transactions. Machine learning is another technique as far as application field. The likelihood of fraud transactions can be anticipated after ATM card transactions when this technique is utilized into bank ATM card fraud detection framework. To keep banks safe from incredible losses and decrease risks, the arrangement of anti-fraud strategies can be accepted. The motivation behind the investigation was taken a few issues contrasted than we had an alternate misclassification cost.

An effective financial statement fraud detection model was introduced by Jan [25] for manageable progress of financial souks confirmation obtained from Taiwan. This model actualizes the various data mining strategies. To screen out critical factors, an artificial neural network (ANN) and a support vector machine (SVM) schemes are masterminded in the main stage. For classification purposes, the four sorts of decision trees, for example, chi-square automatic interaction detector (CHAID), classification and

regression tree (CART), quick unbiased efficient statistical tree (QUEST) and C5.0 are created in the second stage. To build an exceedingly accurate model, both non-financial and financial factors are chosen to recognize fraudulent financial detailing.

Financial fraud detection was produced by Macro *et al* [26] by thinking about human conduct. This article presents detecting fraud, a hypothetical structure that permits to order and synopsis a gathering of individuals inside a financial organization who submit extortion, upheld by the fraud triangle hypothesis. Misrepresentation discovers process in the techniques for a constant audit that will be in dependable of social occasion data of specialists introduced in user apparatus. In view of semantic methods, the data is connected by means of the accumulation of expressions by the clients behind the investigation for later being transferred to a storehouse for later examination. In the reduction of instances of financial fraud, this examination contributes to the cybersecurity field.

Credit card fraud detection was expected by Kuldeep *et al* [27] by utilizing lion's share voting and AdaBoost techniques. To detect credit card fraud, the machine learning method is utilized in this article. At the outset, the standard models are utilized. After that, the hybrid strategies are connected which incorporates the majority voting and AdaBoost procedures. A freely accessible credit card data set is utilized to survey the model adequacy. A genuine credit card data set is investigated which are gotten from the financial institution. Besides, the noise is added to the data samples and also evaluates the toughness of the procedures. In distinguishing fraud cases, the reproduction results demonstrate that the dominant part voting framework gives the better accuracy contrasted with others in the credit card field.

A mining corporate yearly report was depicted by Hajek [28] for detection of financial fraud in relative investigation of machine learning structures. By deriving the consolidating explicit features, this examination recognizes the upgraded fraud detection from financial data and business annual reports. This work needs to finish up a low frequency of harmful words in yearly files it might show non-fraudulent businesses. From freely accessible financial proclamations and predictions of revenues and earnings, it is basic to utilize separated data to recognize the fraudulent firms.

Anticipating fraudulent financial explanations was examined by Stamatis *et al* [29] by utilizing active learning techniques. This work analyses the effectiveness of AL hypothesis over detection of FFS issue contrasted and managed learning techniques. In the relative field, the analog upgrade of the anti-fraud frameworks ought to be utilized to empowering the machine learning instruments. Because of the powerlessness, the plan to gather dependable datasets that assign the equivalent proportions of a firm. It has gone with fraud activities; approaches misuse the presence of few marked events for finding valuable designs from a pool of unlabelled data could be checked extremely capable.

On big data based misrepresentation detection plot was exhibited by Chen *et al* [30] for financial explanations of business gatherings. In finance and economics, this investigation deliberates the several features. Moreover, estimate the big data to upgrade the fraud detection conspire to precisely anticipate the financial statement misrepresentation for financial reports of commercial gatherings. Moreover, improving the investment decision making and diminishing speculation hazards and losses help creditors and investors.

A main impact of social media for financial fraud disclosure was depicted by Dong *et al* [31] in text mining based investigation. In Internet-based life, this examination is used to leverage huge measure of user-produced substance for fraud detection. A logical text structure is framed to detect amorphous internet-based life substances onto words weights, topics, and emotion features. Most of the web-based social networking highlights, proceeding to the fraud which covers the fake and matched identical firms.

For 10-fold cross validation, it gives the average accuracy of 81.43%. For financial fraud revelation, the scheme is prominent aftereffect of social media substance.

Financial fraud detection was presented by Gupta and Nasib [32] by utilizing text mining. For both non-fraud and fraud associations, fraud detection shows a given collection of financial proclamations. In financial proclamations, these two associations are trailed by pre-processing which incorporates a lexical examination of present content. From that point onward, a sack of words plot has been chosen for expelling data hidden in the text which results in vector spaces for both non-fraudulent and fraudulent associations.

Review of the uses of text mining was produced by Kumar and Ravi [33] in the financial domain. For various uses of text mining, this article gives the cutting edge overview. The text ideas are classified as client relationship management (CRM), stock market and forex rate prediction. Some issues are arising which are dominant in optional and consumer advance features. This investigation can be extremely valuable to researchers in this field; the same number of open issues is featured.

By utilizing phonetics, an investigating financial fraud cases based text mining framework was proposed by Zaki and Babis [34]. In various financial settings, this idea bolsters the market observing surveillance frameworks to allow reprocessing of these learning assets, comprehend and share financial fraud logic activities and update the awareness of the procedures of financial fraud. From SEC process, a semantic-based text mining strategy is built up to capture the data and comprehend the unique market manipulation assortments. It helps the examination of fraud during the time spent examination.

An innovative scheme was displayed by Glancy and Yadav [35] for financial statement fraud detection. On textual data, CFDM (computational fraud detection model) utilizes quantifiable technique. For fraud detection, it incorporates a few techniques that utilizes all of the data contained in the textual data in a general sense. In low and high computer-mediated communication (CMC), existing work conveys establishment for perceiving deception. CFDM offers a systematic procedure which prospective for robotics. Imitation annoyed from 10-K filings based on the administration negotiations and investigation. From non-fraudulent scheme, it could discriminate the deceitful filings. CFDM act as a selection device when misleading is distrusted.

Reproduction tried on the MD&A (Management's Discussion and Analysis) from 10-K filings and it could distinguish fraudulent filings from a non-fraudulent scheme. CFDM can help as a screening tool where misleading is suspected.

### **3. Problem Statement**

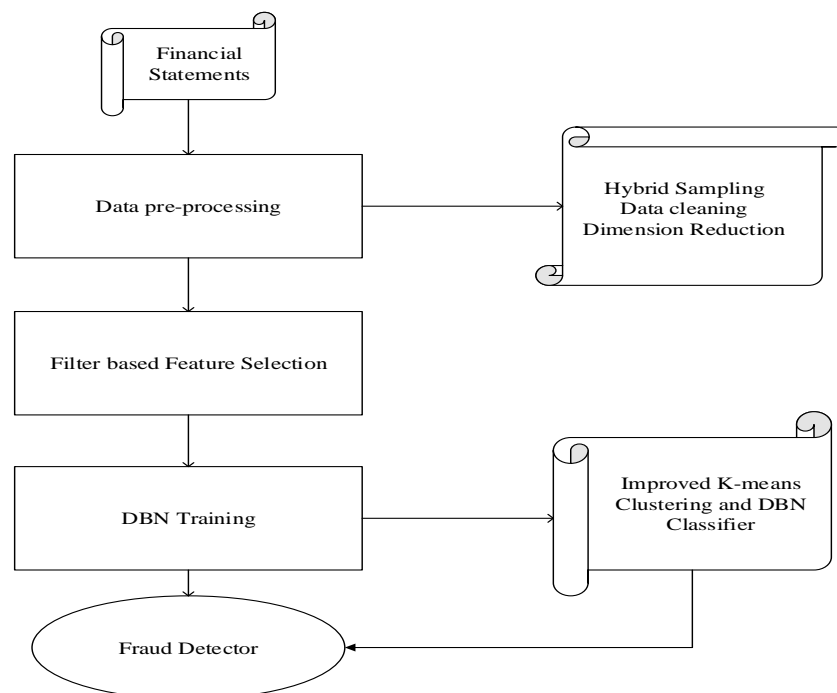
To provide decision support to stakeholder, different data mining strategies have been utilized on the grounds. The financial statement fraud is a troublesome issue for government and community regulator. Public companies make fraudulent financial data just not for the purpose of critical financial loss to the wide investors but also it demotivates the financial market. During the general auditing, the ordinary auditing rehearses that may principally concentrate on factual examination. In financial statements, most financial fraud detection looks into limit their examination just too numerical data. From authentic data, fraudulent financial data could scarcely be recognized because of deliberate covering as well as accounting shenanigans. The common absence of field for financial fraud statement can be hard to diagnose, the rarity in which it happens, and the way that it is typically dedicated by proficient individuals inside the industry who are equipped for marketing their deceit.

### **4. Methodology: A Text Mining Approach**

Most of the companies release the financial statement documents that explain all details about their loans, profits, expenses, and income. From management, it can additionally include some comments based on the business performance. In the future, the predictable problems may arise. The different financial statements provide the pictorial explanation of each company status which is obtained from the company. The stock prices of the successful company are indicated, by various financial statements from which they are identifying whether it is applicable for loans or not.

In this paper, we propose financial statement fraud detection by using effective feature selection and classification technique. This method consists of four stages: 1) pre-processing, 2) sampling, 3) feature selection, and 4) clustering and classification.

In the pre-processing process, data correlation analysis and data cleaning process which cleans the noise data are performed. Also, data transformation, integration, and reduction are included in this process. The following process is the sampling process which evaluates dataset with various ratios for verification through a hybrid sampling method. Feature selection is an effective scheme which used for solving the problems of data mining and machine learning methods. Three kinds of feature selection schemes are available such as filter method, wrapper method, and an embedded method. The wrapper system to assess the selected features based on the analytical presentation of a predefined learning procedure. The filtering scheme to estimate the significance of characteristics depends on the exact features of data and this method is independent in any learning algorithms. With the model learning, an embedded method is to embed the feature selection and it is a mutual technique between the wrapper and filter methods. By using filter-based learning scheme, a feature selection procedure has accomplished. Filter is a method of eliminating the characteristic of low score after score of characteristic based on the score given by evaluation condition. The probability of occurrence of over fitting is relatively low and it is good for large-scale data. Evaluate the importance of characteristics based on reliable data features that are independent of learning algorithms.



**Fig.1:** Financial fraud detection

Clustering process performs using improved K-means after a process of feature selection and this result will be used as training and validation set of the classification process (Fig.1). The optimization process minimizes the objective function (Error Rate) inside the deep belief network. So the optimized deep learning can enhance the fraud detection accuracy.

Financial statement fraud detection can be tested on the datasets such as COMPUSTAT, PagSeguro, Athens stock exchange and 1610 firm-year samples from U.S. listed companies and on any other datasets available in near future. In these datasets, 70% for training and 30% testing can be used.

The performance of the proposed method can be analysed in the form of statistical measures such as recall, sensitivity, F-measure, specificity, and classification accuracy. The simulation outcomes will be evaluated and compared with the earlier classification technique.

## 5. Conclusion

The classification process uses optimized deep belief network (DBN) to get the better accuracy. Moreover, it leads the lower computational cost. A DBN generally have three layers of RBMs (Restricted Boltzmann Machines), the first layer of RBMs takes the input as the visible layer, and the second layer takes the hidden layer as the visible layer. The hidden layer of the last RBM takes the output of DBN as visible layer which is considered as one unit. The above classification processes include the methods:

- I. To analyse the several state-of-art techniques related to financial fraud detection.
- II. To solve the conflicts among the different financial statement fraud.
- III. To analyse the various unsupervised learning algorithms for clustering process of financial statements.
- IV. To get the better classification accuracy for financial statements fraud by extracting the several optimal features under the feature selection process.

The performance of the proposed method can be analysed in the form of statistical measures such as recall, sensitivity, F-measure, specificity, and classification accuracy. The simulation outcome may be evaluated and compared with the earlier classification techniques.

## 6. References

- [1] Sumathy, K. L., and M. Chidambaram. "Text mining: concepts, applications, tools, and issues-an overview." *International Journal of Computer Applications* 80, no. 4 (2013).
- [2] Aggarwal, Charu C., and Haixun Wang. "Text mining in social networks." In *Social network data analytics*, pp. 353-378. Springer, Boston, MA, 2011.
- [3] Mostafa, Mohamed M. "More than words: Social networks' text mining for consumer brand sentiments." *Expert Systems with Applications* 40, no. 10 (2013): 4241-4251.
- [4] Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Moshe Fresko. "Mine your own business: Market-structure surveillance through text mining." *Marketing Science* 31, no. 3 (2012): 521-543.
- [5] Fuller, Christie M., David P. Biros, and Dursun Delen. "An investigation of data and text mining methods for real-world deception detection." *Expert Systems with Applications* 38, no. 7 (2011): 8392-8398.



- [6] Othman, Rohana, Nooraslinda Abdul Aris, Ainun Mardiyah, Norhasliza Zainan, and Noralina Md Amin. "Fraud detection and prevention methods in the Malaysian public sector: Accountants' and internal auditors' perceptions." *Procedia Economics and Finance* 28 (2015): 59-67.
- [7] Dong, Wei, Shaoyi Liao, and Liang Liang. "Financial Statement Fraud Detection using Text Mining: A Systemic Functional Linguistics Theory Perspective." In *PACIS*, p. 188. 2016.
- [8] Fu, Kang, Dawei Cheng, Yi Tu, and Liqing Zhang. "Credit card fraud detection using convolutional neural networks." In *International Conference on Neural Information Processing*, pp. 483-490. Springer, Cham, 2016.
- [9] Rawte, Vipula, and G. Anuradha. "Fraud detection in health insurance using data mining techniques." In *Communication, Information & Computing Technology (ICCICT), 2015 International Conference on*, pp. 1-5. IEEE, 2015.
- [10] Dilla, William N., and Robyn L. Raschke. "Data visualization for fraud detection: Practice implications and a call for future research." *International Journal of Accounting Information Systems* 16 (2015): 1-22.
- [11] Kanapickienė, Rasa, and Živilė Grundienė. "The model of fraud detection in financial statements by means of financial ratios." *Procedia-Social and Behavioral Sciences* 213 (2015): 321-327.
- [12] West, Jarrod, and Maumita Bhattacharya. "Some Experimental Issues in Financial Fraud Mining." In *ICCS*, pp. 1734-1744. 2016.
- [13] Kim, Yeonkook J., Bok Baik, and Sungzoon Cho. "Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning." *Expert systems with applications* 62 (2016): 32-43.
- [14] Olszewski, Dominik. "Fraud detection using self-organizing map visualizing the user profiles." *Knowledge-Based Systems* 70 (2014): 324-334.
- [15] Albrecht, Chad, Daniel Holland, Ricardo Malagueño, Simon Dolan, and Shay Tzafrir. "The role of power in financial statement fraud schemes." *Journal of Business Ethics* 131, no. 4 (2015): 803-813.
- [16] West, Jarrod, Maumita Bhattacharya, and Rafiqul Islam. "Intelligent financial fraud detection practices: an investigation." In *International Conference on Security and Privacy in Communication Systems*, pp. 186-203. Springer, Cham, 2014.
- [17] Lin, Chi-Chen, An-An Chiu, Shaio Yan Huang, and David C. Yen. "Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments." *Knowledge-Based Systems* 89 (2015): 459-470.
- [18] McMahon, Richard, Diana Pence, Linda Bressler, and Martin S. Bressler. "New Tactics in Fighting Financial Crimes: Moving Beyond the Fraud Triangle." *Journal of Legal, Ethical & Regulatory Issues* 19, no. 1 (2016).
- [19] Wong, Shirley, and Sitalakshmi Venkatraman. "Financial accounting fraud detection using business intelligence." *Asian Economic and Financial Review* 5, no. 11 (2015): 1187.
- [20] Throckmorton, Chandra S., William J. Mayew, Mohan Venkatachalam, and Leslie M. Collins. "Financial fraud detection using vocal, linguistic and financial cues." *Decision Support Systems* 74 (2015): 78-87.

- [21] Huang, Dongxu, Dejun Mu, Libin Yang, and Xiaoyan Cai. "CoDetect: Financial Fraud Detection with Anomaly Feature Detection." *IEEE Access* 6 (2018): 19161-19174.
- [22] Yao, Jianrong, Jie Zhang, and Lu Wang. "A financial statement fraud detection model based on hybrid data mining methods." In *2018 International Conference on Artificial Intelligence and Big Data (ICAIBD)*, pp. 57-61. IEEE, 2018.
- [23] Yee, Ong Shu, Saravanan Sagadevan, and Nurul Hashimah Ahamed Hassain Malim. "Credit Card Fraud Detection Using Machine Learning as Data Mining Technique." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 10, no. 1-4 (2018): 23-27.
- [24] Prakash, B., G. Venu Madhava Murthy, P. Ashok, B. Pavan Prithvi, and S. Sai Harsha Kira. "ATM Card Fraud Detection System Using Machine Learning Techniques." (2018).
- [25] Jan, Chyan-long. "An Effective Financial Statements Fraud Detection Model for the Sustainable Development of Financial Markets: Evidence from Taiwan." *Sustainability* 10, no. 2 (2018): 513.
- [26] Sánchez, Marco, Jenny Torres, Patricio Zambrano, and Pamela Flores. "FraudFind: Financial fraud detection by analyzing human behavior." In *Computing and Communication Workshop and Conference (CCWC)*, 2018 IEEE 8th Annual, pp. 281-286. IEEE, 2018.
- [27] Randhawa, Kuldeep, Chu Kiong Loo, Manjeevan Seera, CHEE PENG Lim, and Asoke K. Nandi. "Credit card fraud detection using AdaBoost and majority voting." *IEEE ACCESS* 6 (2018): 14277-14284.
- [28] Hajek, Petr, and Roberto Henriques. "Mining corporate annual reports for intelligent detection of financial statement fraud—a comparative study of machine learning methods." *Knowledge-Based Systems* 128 (2017): 139-152.
- [29] Karlos, Stamatis, Georgios Kostopoulos, Sotiris Kotsiantis, and Vassilis Tampakas. "Using Active Learning Methods for Predicting Fraudulent Financial Statements." In *International Conference on Engineering Applications of Neural Networks*, pp. 351-362. Springer, Cham, 2017.
- [30] Chen, Yuh-Jen, and Chun-Han Wu. "On Big Data-Based Fraud Detection Method for Financial Statements of Business Groups." In *Advanced Applied Informatics (IIAI-AAI)*, 2017 6th IIAI International Congress on, pp. 986-987. IEEE, 2017.
- [31] Dong, Wei, Shaoyi Liao, Yujing Xu, and Xiaoqian Feng. "Leading effect of social media for financial fraud disclosure: A text mining based analytics." (2016).
- [32] Gupta, Rajan, and Nasib Singh Gill. "Financial statement fraud detection using text mining." *Editorial Preface* 3, no. 12 (2012): 189-191.
- [33] Kumar, B. Shravan, and Vadlamani Ravi. "A survey of the applications of text mining in financial domain." *Knowledge-Based Systems* 114 (2016): 128-147.
- [34] Zaki, Mohamed, and Babis Theodoulidis. "Analyzing financial fraud cases using a linguistics-based text mining approach." (2013).
- [35] Glancy, Fletcher H., and Surya B. Yadav. "A computational model for financial reporting fraud detection." *Decision Support Systems* 50, no. 3 (2011): 595-601.