Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2024, 11(5):84-93

Research Article ISSN: 2394 - 658X

Real-Time Analytics for Banking: Enhancing Decision-Making and Fraud Detection

Pranay Mungara

Engineering Managerm, Business Intelligence, Virtusa Corporation, Wesley Chapel, FL, USA
**Email: pranay.mungara@gmail.com*

ABSTRACT

Innovative and quick detection procedures are necessary as financial institutions deal with the increasing prevalence of fraudulent operations. This assessment highlights the revolutionary possibilities of Big Data Analytics and how it may play a key role in the continuous battle against fraud. Using predictive modelling and machine learning algorithms, the piece examines a variety of data sources including data on user behaviour and transactions, as well as external data gathered from sites like social media in order to spot outliers and assess risk. The use of real-time processing is becoming more and more critical for effective and quick fraud detection. In this article, we examine the rising value of data and the impact of big data on the banking business. Banks are reshaping their operations, client communications, and long-term planning with the help of big data analytics. Following an analysis of big data with an emphasis on its volume, velocity, and diversity, a comprehensive review of its beginnings in the banking industry's marketing history, customers and communications, social media interaction, and other areas is conducted. Next, the paper delves into how it will affect banking operations. The article goes into detail about how banks are using big data to manage risk better, improve client services, increase operational efficiency, and uncover fraudulent behaviour. Big data is also adding to the banking industry's huge database. The essay concludes by discussing the difficulties that banks have when dealing with big data, including concerns over data privacy and security, and by presenting case studies that show how banks can profit from big data analytics.

Keywords: Banking, Decision-making, Fraud detection

---

1. INTRODUCTION

Fraud in the financial sector, in all of its myriad forms, has arisen as a ubiquitous concern in the contemporary environment of international business. The level of sophistication and adaptability of fraudulent schemes is constantly increasing, which continues to provide issues that need for countermeasures that are both new and dynamic. Identity theft, fraudulent use of payment cards, and cybercrime are all examples of some of the behaviours that fall under the umbrella of financial fraud, which has become increasingly widespread and sophisticated. Both opportunities and challenges have arisen as a result of the incorporation of technology into financial transactions. This has resulted in the creation of an environment in which criminals take advantage of weaknesses in order to make illegal gains. A thorough comprehension of the scope and dynamics of modern financial fraud is crucial for developing effective prevention strategies. The methods that are used to commit financial fraud are dynamic and adaptable, and they are always changing in order to take advantage of vulnerabilities in technological, social, and economic systems (Smith & Johnson, 2020) [1].

It is possible for criminals to evade conventional security measures by employing novel strategies, which might range from standard kinds of fraud to complicated cyberattacks for example. Given the nature of the situation, it

is necessary to take a proactive and adaptable strategy to detection and prevention. The consequences of financial fraud extend beyond monetary losses and include damage to reputation as well as the erosion of trust. These implications have an impact on both businesses and financial institutions. The majority of direct financial losses are borne by financial institutions, while businesses are need to deal with disruptions in their operations as well as the possibility of legal repercussions. Because of the linked nature of the global financial system, the ripple effects of fraud are amplified, which highlights the vital need for robust detection measures. It is of the utmost importance to undertake early detection and preventive measures in order to mitigate the cascade impacts of financial fraud (Othman et al, 2020) [2].

Intervention in a timely manner not only reduces the amount of money that is lost, but it also protects the integrity of the financial systems and provides protection for the interests of both individuals and businesses. The proactive identification of fraudulent actions is dependent on the utilisation of cutting-edge technology, such as Big Data Analytics, which are capable of performing real-time analysis on several massive datasets. An analysis and evaluation of the crucial function that Big Data Analytics plays in the landscape of financial fraud detection is going to be the primary subject of this article. Big Data Analytics is a technique that utilises sophisticated algorithms to analyse large datasets in order to identify trends, abnormalities, and potential dangers. It is essential for stakeholders who are looking for effective countermeasures to have a solid understanding of how this technology contributes to the detection and prevention of financial crime. Detecting fraudulent activity becomes increasingly difficult as the financial environment continues to change (Patel, 2023) [3].

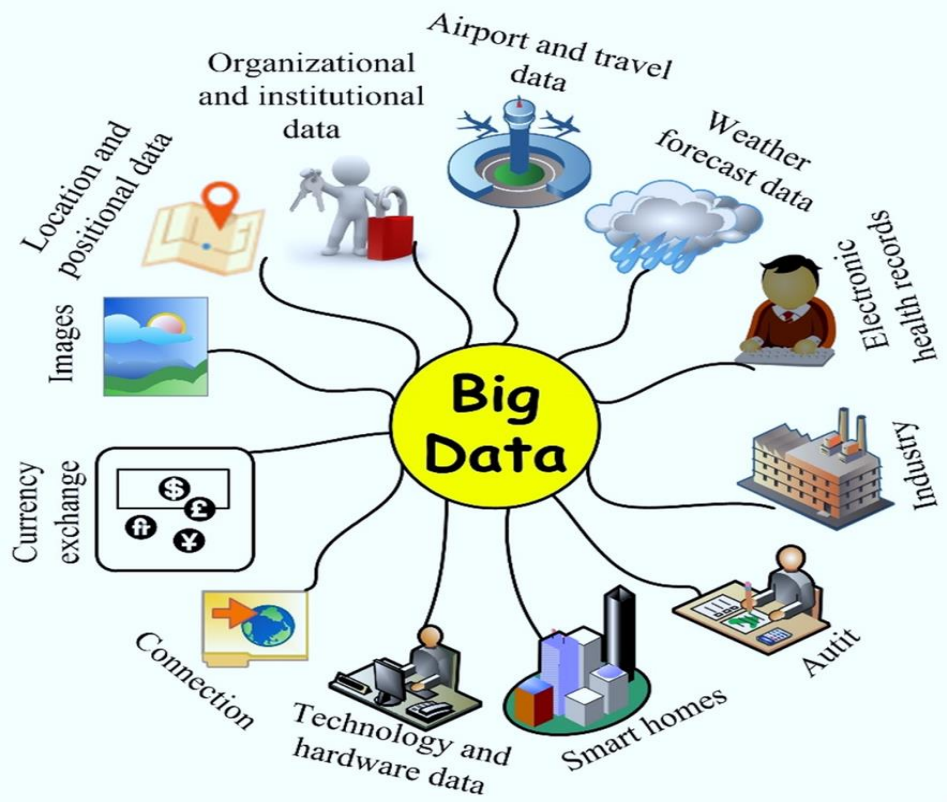

Figure 1: Big data in real-time

Big Data is characterised by four basic features, which are commonly referred to as the 4Vs. These characteristics encompass the distinct difficulties and opportunities that are presented by large-scale data analytics. Big Data Analytics is concerned with dealing with enormous amounts of data, which are frequently generated in real time. This enormous scale necessitates the implementation of scalable infrastructure and processing capabilities that are efficient. Monitoring the data generation, processing, and analysis rates is crucial. To obtain timely insights, it is crucial to do processing in real time or near real time, according to Chen et al. (2023), [4]. There are many different formats that Big Data can be found in, such as structured, unstructured, and semi-structured data varieties. This variability necessitates the utilisation of adaptable tools and methods that are able to manage a wide variety of data kinds.

2. LITERATURE REVIEW

The banking sector has been profoundly impacted by the advent of information technology (IT). Credit card and online net banking transactions currently account for the vast bulk of banking system transactions (Jiang and Broby, 2021) [5].

These transactions all create extra vulnerabilities. The number of hackers targeting institutions that have massive amounts of customer information has been growing. Consequently, when it comes to company cyber security, financial institutions have been at the front. The field of cyber security has grown substantially within the last 13 years.

A valuation of 170.4 billion dollars is anticipated for the market in the year 2022. It is anticipated that the cost of cybercrime will increase by fifteen percent annually over the following three years, eventually surpassing ten and a half trillion dollars annually by the year 2025 (Morgan, 2020) [6].

There is a huge concern in the banking business regarding cyber fraud that involves credit cards, which results in annual costs of billions of dollars. One of the most important priorities for the banking business is to increase cyber security protection. Numerous methods have been created for the purpose of monitoring and identifying instances of cyber fraud using credit cards. But because threats are always evolving, financial institutions must have state-of-the-art cyber fraud management systems (Btoush et al., 2021). [7].

Cyber fraud involving credit cards has been on the rise due to the phenomenal growth in the use of credit cards and other forms of online payment in recent years. When it comes to credit cards, there are a few distinct forms of cyber fraud. To begin with, there is the classic case of a credit card being stolen. The second type of cyber fraud involves the theft of sensitive information pertaining to credit card details. If the cardholder doesn't authorise the entry of their credit card information during an online purchase, extra fraudulent activity might happen (Al Smadi & Min, 2020; Trivedi et al., 2020) [8, 9].

Researchers in the area of machine learning (ML) have taken an interest in the challenging problem of detecting fraudulent purchases made with credit cards. The datasets associated with credit cards exhibit a significant degree of skewness. Lots of algorithms fail to distinguish between items in minority classes when dealing with extremely skewed datasets. Online fraud detection systems need to be able to react quickly if they are to be efficiently deployed. Another major issue is how new attack tactics change the conditional distribution of the data over time (Benchaji, Douzi, and El Ouahidi, 2021) [10].

Al Rubaie (2021) [11] states that finding cyber fraud in credit card transactions accurately requires overcoming a lot of challenges. Here we discuss a number of challenges, including real-time detection, vast volumes of data that are imbalanced or incorrectly classified, and frequent changes in the kind of transaction.

Detecting cyber fraud is becoming an increasingly important field as the technology that is currently available continues to advance. Cyber credit card fraud is also expanding at a quick pace. The traditional approaches that have been used in the past to handle this issue are no longer efficient. Using the traditional method, domain experts in the field of cyber fraud are responsible for composing the algorithms, which are governed by stringent standards. Additionally, in order to counteract cyber fraud, a proactive approach is required to be implemented. Every industry is trying to adopt machine learning-based solutions since they are fast, efficient, and widely used (Priya & Saradha, 2021) [12]. Results show that ML and DL methods work well for this research field. DL has received the lion's share of media attention and has been the most effective countermeasure against cyber assaults as of late. Its exceptional value in this sector stems, in part, from its capacity to process massive datasets, reduce the probability of overfitting, and reveal hidden patterns of fraud.

Artificial intelligence (DL) approaches have been utilised throughout the course of the past several years to identify new fraudulent trends and to enable systems to respond in a flexible manner to complicated data patterns. In order to present the most recent and pertinent information on the subject, we have decided to concentrate on the most recent research that has been conducted between the years 2019 and 2021. This is due to the fact that the popularity of DL has increased over this time period.

3. BIG DATA APPLICATIONS THE BANKING SECTOR

The process of extracting useful information from historical data was the first step in the development of business intelligence. According to Dicuonzo et al. (2019), the extraction procedure was accomplished through the building of data storage archives. These archives were made to efficiently manage and evaluate massive amounts of data quickly. There are thirteen of them. Internationalisation and technological advancements have

made corporate management more complicated. This is because company executives are looking for more backing for the decision-making policies they employ. Decisions based on data are becoming more important and valuable as a result. There are two main currents in big data analytics that companies are focusing on. The first topic that will be covered is businesses that use big data analytics to find new opportunities. The optimisation of internal processes and the improvement of already-existing products or services are two other applications (Hung et al., 2020). We have [14]. Conversely, emerging companies use big data analytics extensively to create innovative products and services. Also, when it comes to using analytical data to make judgements, the banking business is a good place to start. "Big data" refers to "large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information" (Hassani et al., 2018a).

This definition is based on the fact that big data is a term that has been used to describe the phenomenon. It is [15]. The availability of such a tool has provided businesses all over the world with the opportunity to study the vast volumes of data that they possess in search of abnormalities or information that is foreign to them, which can be useful in the process of decision-making. Prior research has determined that there are five characteristics that are associated with big data. Data volume (the amount of data), data generation speed, data value, data accuracy, and data format variability (structured, semi-structured, and unstructured data) are all relevant features (Dicuonzo et al., 2019) [16].

With the help of big data, financial institutions are able to provide their customers with a wide range of customised goods, which in turn boosts their overall performance rate. In addition, Al-Dmour et al. (2021) [17] came to the conclusion in their research that the implementation of BDA will have a favourable impact on the functioning of economic institutions. According to the findings of a previous research group, the functions of big data in the banking business may be broken down into three categories: "customer relationship management (CRM), fraud detection and prevention, and risk management and investment banking."

One of the most effective financial technology tools that has been introduced to the market to support company sustainability objectives is big data analytics. To be more specific, BDA makes it possible for enterprises, particularly banks, to enhance their internally functioning operations. By analysing possible risk predictors, banks may, for instance, find key components that could increase their client retention rate. In the present market, when there is an excess of competition, this would guarantee a company's continuous existence. There are several benefits of using fintech, which includes BDA, in commercial banks, according to Wang et al. (2021) [18]. Better business models, lower operational costs, more efficient services, better risk control, and more competitiveness are all benefits. Regardless, before businesses and financial institutions decide to use BDA, they need to set up the infrastructure that will drive their success. This organisation is supported by BDA hardware equipment and top-notch big data software.

The audience will be equipped with the knowledge they need to understand and benefit from the outcomes of big data analysis thanks to the study's findings. Big data has been the subject of extensive literature reviews for nearly four decades, from 1960 to 2020. Utilising bibliometric analysis in conjunction with network analysis allowed for this evaluation to be conducted.

The findings of this study provide insights on the quality of contributions made by researchers as well as the most recent advancements made towards this subject in scientific publications. In addition, bibliometric analysis has been of use in determining the most important countries, authors, and research clusters that are associated with this particular field of study (Nobanee, 2020) [19].

This bibliographic review analyses the data using a variety of approaches, each of which has been utilised. Scopus is used as a search engine to filter papers that were published between the years 1960 and 2020 for this study. After the year 2005, the majority of the articles that are pertinent were published. The citation rate, which indicates the number of times prior research has referred to a particular study, has been used to categorise the research papers. Given that the study has earned the highest citation rate, it is clear that academics have given it the most significant attention due to the fact that it contains valuable content. The data acquired for this study is displayed using various tables and figures according to the citation rate. This study (Nobanee, 2021) [20] also provides a visual representation of the documents, authors, co-authors, citations, connection strengths, application organisation, and nations of publications that were extracted.

4. CYBER FRAUD DETECTION USING MACHINE AND DEEP LEARNING

4.1 Unsupervised techniques

Clustering is the technique of classifying occurrences that are similar into groupings that are identical to one another. A significantly smaller amount of clustering strategies were employed in comparison to classification methods in the paper that was examined. Modelling the probability distribution over sequences of observations is accomplished through the utilisation of the hidden Markov model. Both hidden states and outputs that can be observed make up this system. There have been seven articles that have utilised HMM. For the purpose of detecting cyber attacks, Das et al. (2020) [21] utilised an HMM model. Not only do the results reveal that the suggested system performs exceptionally well, but they also highlight the benefits of learning cardholders' spending patterns.

Singh et al. (2021) [22] proposed a method to determine the spending profile of cardholders, after which they made an effort to discover the observation symbols. These observation symbols will be of assistance in determining the initial estimate of the model parameters. Consequently, HMM is able to determine if the transaction is fake or genuine. In conjunction with HMM and density-based spatial clustering of applications and noise, SMOTE was employed. Compared to the previous model, this new one (SMOTE+DBSCAN+HMM) performed relatively better for all of the different hidden states.

A total of seven articles have utilised the K-means algorithm. Data clustering can be accomplished by the use of the K-means algorithm, which is a non-hierarchical method. The algorithm employs a straightforward approach. Consequently, K-means is able to categorise a given dataset into a predetermined number of groups, also known as K-clusters. The K-mean algorithm was utilised in conjunction with the back-propagation neural network (BPNN) in Abdulsalami et al. (2019) [23]. BPNN and K-means are both used to detect fraudulent credit card transactions; nevertheless, the results indicate that there is a significant difference between the two methods. Comparing the BPNN model to the K-means model, the BPNN model attained a higher level of accuracy while producing fewer false alarms. In addition, the findings demonstrate that the accuracy of BPNN is 93.1%, whereas the accuracy of K-means is 79.9%.

An unsupervised ensemble is that of an isolation forest. There is no evaluation of regular instances, nor is there any calculation of distance based on points. Instead, an ensemble of DTs is constructed by the Isolation forest. To divide anomalies with the intention of isolating them is the idea behind the isolation forest concept. In the process of generating an ensemble of DTs for a specific data collection, the data points that have the average path length that is the shortest are distinguished as being anomalous. There have been 19 articles that make use of the isolation forest. Meenu et al. (2020) [24] makes use of a brand new Isolation Forest model in order to identify instances of cheating. An methodology that is much superior to other methods of fraud detection is demonstrated by the model, which exhibits the efficiency in detecting fraud, which was seen to be 98.72%.

Vijayakumar et al. (2020) [25] have applied the concept of isolation forest with local outlier factor in order to detect fraudulent activity. Both the isolation forest and the local outlier factor demonstrated a level of accuracy of 99.72% and 99.62% respectively. When conducting business online, the isolation aspect is more easily recognised. Palekar et al. (2020) [26] conducted a study that demonstrated that K-means clustering and (Isolation forest and local outlier factor) may be produced and created on a very large scale for the purpose of detecting fraudulent activity in credit card transactions.

NN stands for unsupervised neural networks learning, which is what self-organizing maps (SOM) are. The SOM is suitable for the purpose of constructing and assessing customer profiles in order to identify fraudulent activity. Two of the articles that were reviewed utilised SOM.

Harwani et al. (2020) [27] utilised SOM and NN in a hybrid strategy to investigate the problem. In comparison to the use of SOM and ANN on their own, the model that was suggested achieved a higher level of accuracy and cost. K-means, K-means clustering utilising principal component analysis (PCA), T-distributed stochastic neighbour embedding (T-SNE), and SOM are the three unsupervised techniques that are presented in Deb, Ghosal, and Bose's (2021) [28] paper. The accuracy of this model for detecting fraudulent activity in credit card transactions was 90%. The findings indicate that K-means clustering used in conjunction with principal component analysis is superior to K-means alone. T-SNE is also superior to PCA since the latter is significantly impacted by outliers, but T-SNE is not.

4.2 Semi-supervised techniques

One method that combines supervised and unsupervised learning is called a hybrid methodology. When determining the best way to represent data, the unsupervised learning attribute is utilised. On the other hand, the supervised learning attribute is applied to study the relationships that are present in the representation before beginning to make predictions. In situations where the data gathering is not balanced, semi-supervised learning is an incredibly effective technique. Each of the studies that are included in this review conducted their study using a semi-supervised methodology. A total of three research utilised semi-supervised methods to identify instances of credit card fraud. The authors Dzakiyullah, Pramuntadi, and Fauziyyah (2021) [29] offer a method that utilises semi-supervised learning in conjunction with AutoEncoders in order to identify fraudulent transactions. An autoencoder was utilised in this proposed model, which was followed by the training of a basic linear classifier in order to assign the data collection to its own category. The T-SNE was also utilised in order to show the fundamental differences between fraudulent and non-fraudulent transactions. Due to the fact that credit card fraud may be easily classified with 0.98%, the results that were achieved in this study are beneficial.

In Pratap and Vijayaraghavulu's (2021) [30] work, semi-supervised algorithms that use majority voting were utilised; in this study, twelve machine learning algorithms were applied respectively. The first step is to make use of the standard models. Second, the addition of AdaBoost and support for majority voting. Based on the results, it appears that the approach of majority voting achieves a high level of accuracy.

4.3 Deep learning

Deep learning (DL) is a subfield of machine learning that employs data to instruct computers on how to carry out certain tasks. One of the most important principles of deep learning is that the performance of our neural networks (NN) continues to improve as we extend them and train them with new data. The primary benefit of deep learning in comparison to regular machine learning is that it performs better on large datasets. The feed forwards neural networks (FNNs), stacked autoencoders (SAE), and convolutional neural networks (CNNs) are the various deep learning algorithms that are utilised in the field of cybersecurity the most frequently. It is the autoencoder (AE) that is responsible for discovering the input-output mapping that exists between the encoding and decoding phases. The input is mapped to the hidden layer by the encoder, and the decoder reconstructs the input by using the hidden layer as the output layer. Both of these processes take place in the same order. Each of the 18 articles included in this review featured AE. 18 of the articles included in this review make reference to AE. An application of the autoencoder model for the detection of cyber fraud is shown in Misra et al. (2020) [31]. A two-stage model that includes an autoencoder that, in the first step, converts the characteristics of the transaction into a feature vector with a lower dimension. These feature vectors are then supplied into a classifier in the succeeding step of the process. A comparison of the suggested model to other models reveals that it performs better.

Dual autoencoders generative adversarial networks (DAEGAN) are utilised in Wu, Cui, and Welsch's (2020) [32] research in order to solve the unbalanced classification problem. For the purpose of autoencoder training, the new model instructs GAN to repeatedly replicate fraudulent transactions. In order to generate two distinct sets of features, the samples are encoded by two autoencoders. This novel model performs better than a number of different categorization techniques. The classification situations that are presented by credit card datasets are unbalanced since the class distributions are excessively skewed. In order to cope with this challenge. The new model proposes using an oversampling strategy that is based on variational automated coding (VAE) in conjunction with deep learning techniques. The findings indicate that the VAE model performs better than both synthetic minority oversampling strategies and conventional distributed neural network (DNN) methods. Furthermore, it displays superior performance in comparison to earlier oversampling methods that were based on GAN models.

4.4 Metaheuristic techniques

In the study of Makolo and Adeboye (2021) [33], a new hybrid model is developed by using the Genetic algorithm and the multivariate normal distribution on a dataset that is incompletely balanced. A comparison was made between the accuracy of the prediction made by DT, ANN, and SVM after they were trained on the same dataset. A remarkable F-score of 93.5% was obtained by the model, in contrast to the scores of 68.5% for ANN,

80.0% for DT, and 84.2% respectively for SVM. The enhanced hybrid system for the prediction of fraudulent credit card activity. It is used to implement the genetic algorithm with RF model optimisation, often known as GAORF. utilising both evolutionary and real-world techniques. An improvement in the accuracy of this model's classification was achieved through the optimisation of RF models. It is possible that this will be of assistance in resolving the issue of insufficient optimisation and convergence of RF algorithms, as well as the issue of a lack of transaction data. The overall number of incorrect classifications was greatly reduced as a result of the model's major improvement.

It is reported in Daliri (2020) [34] that the employment of the harmony search algorithm (HAS) in conjunction with NN can boost the detection of fraudulent activity. In order to optimise the parameters of the ANN, the model makes use of HAS. The NNHS model that has been proposed offers a strategy that is founded on HAS and is capable of accurately predicting the best structure for ANN as well as locating the algorithm that is concealed inside the data. All of the comparisons showed that the best level of accuracy that could be reached was 86%.

4.5 Instance-based learning

This work was carried out by Hussein, Abbas, and Mahdi (2021) [35], which created a fraud detection model that made use of a number of different machine learning algorithms. These algorithms included NB, DR, rules classifier, lazy classifier (IBK, LWL, and KStar), meta classifier, and function classifier. According to the findings, the technique known as the lazy classifier (LMT) offers the highest level of accuracy, with an accuracy of 82.086%.

One hundred forty-four percent of the selected articles made use of the supervised method, as shown in Figure 2. In light of this, the supervised strategy is the one that is utilised the most frequently in the article that was examined. On the other hand, twelve percent of the participants employed unsupervised methods, and twelve percent utilised both supervised and unsupervised methods. A total of two percent of the articles that were examined made use of semi-supervised learning. In addition to that, one percent made use of reinforcement learning. In the years 2019, 2020, and 2021, both supervised and unsupervised learning have been developed and deployed. whereas the use of semi-supervised learning was only three times in the year 2021. In a similar fashion, the year 2021 was the first year when reinforcement learning was implemented. When compared to supervised and unsupervised learning, semi-supervised learning and reinforcement learning were not widely accepted by researchers. This is because of the challenges that they presented. Figure 2 depicts the percentage of supervised, unsupervised, and semi-supervised activities in the overall sample.

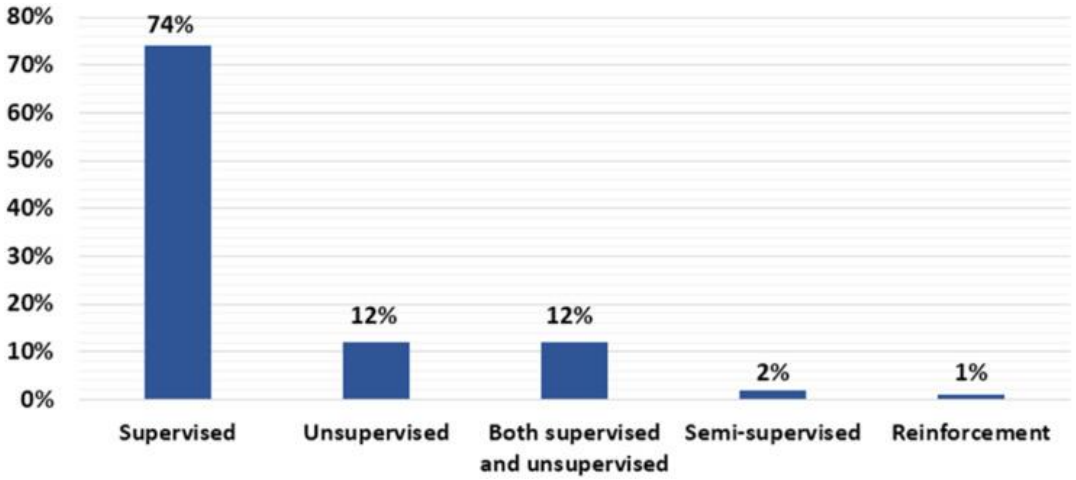

Figure 2: Percentage of supervised, unsupervised, or semi-supervised

5. CONCLUSION

The banking sector has seen substantial transformations due to the advancement of technology. The impact of big data analytics on the financial industry has made it the most rapidly expanding trend. This is useful for banks that deal with massive amounts of data. Has used a bibliographic review to examine key themes within

the realm of big data as they pertain to banks, including risk assessment, financial management, credit risk, customer analysis, bankruptcy prediction, anti-fraud systems, investment, profit, strategic framework management, and competitiveness. More and more nations are getting involved in this field of study to learn more about the advantages and growing significance of banking services.
Based on the citation rate and link strength, these documents have been found and filtered in this research. Because the citation rate indicates that researchers are becoming more aware of this discipline, this technique was effective in this bibliometric study. As a result, researchers are now more willing to shed light on the significance of this topic by conducting important research. Using a top-down approach, this article has charted the landscape of big data in banking and identified emerging patterns.
Credit card cyber fraud detection using ML/DL approaches was the focus of this review. From the vantage points of learning-based fraud detection, ML/DL performance estimation, and ML/DL technique type, we analysed ML/DL models. Articles published in 2019, 2020, and 2021 were the primary focus of the research. We combed through 181 articles to find answers to the four research topics that guided this investigation. Our review has covered all the bases, including an examination of ML/DL methods and their role in identifying cyber fraud on credit cards, as well as suggestions for picking the best methods.

REFERENCES

[1]. Smith, J., & Johnson, A. (2020). The Dynamics of Financial Fraud in the Digital Age. Journal of Financial Crime, 27(2), 527-543.

[2]. Othman, Z., Nordin, M. F. F., & Sadiq, M. (2020). GST fraud prevention to ensure business sustainability: a Malaysian case study. Journal of Asian Business and Economic Studies, 27(3), 245-265.

[3]. Patel, K. (2023). Credit Card Analytics: A Review of Fraud Detection and Risk Assessment Techniques. International Journal of Computer Trends and Technology, 71(10), 69-79.

[4]. Chen, W., Milosevic, Z., Rabhi, F. A., & Berry, A. (2023). Real-time analytics: concepts, architectures and ML/AI considerations. IEEE Access

[5]. Jiang & Broby (2021) Jiang C, Broby D. Mitigating cybersecurity challenges in the financial sector with artificial intelligence. 2021. https://pure.ulster.ac.uk/files/98691946/Jiang_Broby_CeFRI_2021_Mitigating_cybersecurity_challenges_in_the_financial_sector_with_Artificial_Intelligence.pdf

[6]. Morgan (2020) Morgan S. Cyberwarfare in the C-suite. 2020. [2 June 2021]. https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/ Cybercrime Magazine.

[7]. Btoush et al. (2021) Btoush E, Zhou X, Gururaian R, Chan KC, Tao X. A survey on credit card fraud detection techniques in banking industry for cyber security. 2021 8th International Conference on Behavioral and Social Computing (BESC); Piscataway: IEEE; 2021. pp. 1–7

[8]. Al Smadi & Min (2020) Al Smadi B, Min M. A critical review of credit card fraud detection techniques. 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON); Piscataway: IEEE; 2020. pp. 732–736.

[9]. Trivedi et al. (2020) Trivedi NK, Simaiya S, Lilhore UK, Sharma SK. An efficient credit card fraud detection model based on machine learning methods. International Journal of Advanced Science and Technology. 2020;29(5):3414–3424

[10]. Benchaji, Douzi & El Ouahidi (2021) Benchaji I, Douzi S, El Ouahidi B. Credit card fraud detection model based on LSTM recurrent neural networks. Journal of Advances in Information Technology. 2021;12(2):113–118. doi: 10.12720/jait.12.2.113-118.

[11]. Al Rubaie (2021) Al Rubaie EMH. Improvement in credit card fraud detection using ensemble classification technique and user data. International Journal of Nonlinear Analysis and Applications. 2021;12(2):1255–1265. doi: 10.22075/IJNAA.2021.5228.

[12]. Priya & Saradha (2021) Priya GJ, Saradha S. Fraud detection and prevention using machine learning algorithms: a review. 2021 7th International Conference on Electrical Energy Systems (ICEES); Piscataway: IEEE; 2021. pp. 564–568.

[13]. Dicuonzo G., Galeone G., Zappimbulso E., Dell'Atti V. (2019). Risk management 4.0: The role of big data analytics in the bank sector. *International Journal of Economics and Financial Issues*, 9(6), 40–47.

[14]. Hung J. L., He W., Shen J. (2020). Big data analytics for supply chain relationship in banking. *Industrial Marketing Management*, 86, 144–153.

[15]. Hassani H., Huang X., Silva E. (2018a). Digitalisation and big data mining in banking. *Big Data and Cognitive Computing*, 2(3), 18.

[16]. Dicuonzo G., Galeone G., Zappimbulso E., Dell'Atti V. (2019). Risk management 4.0: The role of big data analytics in the bank sector. *International Journal of Economics and Financial Issues*, 9(6), 40–47.

[17]. Al-Dmour H., Saad N., Basheer Amin E., Al-Dmour R., Al-Dmour A. (2021). The influence of the practices of big data analytics applications on bank performance: Filed study. *VINE Journal of Information and Knowledge Management Systems*. Advance online publication. https://doi.org/10.1108/VJIKMS-08-2020-0151

[18]. Wang Y., Xiuping S., Zhang Q. (2021). Can fintech improve the efficiency of commercial banks?—An analysis based on big data. *Research in International Business and Finance*, 55, 1–28.

[19]. Nobanee H. (2020). Big data in business: A bibliometric analysis of relevant literature. *Big Data*, 8(6), 459–463.

[20]. Nobanee H. (2021). A bibliometric review of big data in finance. *Big Data*, 9(2), 73–78.

[21]. Das et al. (2020) Das TA, Lagade KC, Girase MP, Patole R. Credit card fraud detection system using data mining. Artificial & Computational Intelligence. 2020;3(3)

[22]. Singh et al. (2021) Singh GK, Bhayye A, Dhamnaskar S, Patil S, Phulari SV. Credit card fraud detection using isolation forest. International Journal of Recent Advances in Multidisciplinary Topics. 2021;2(6):118–119.

[23]. Abdulsalami et al. (2019) Abdulsalami BA, Kolawole AA, Ogunrinde MA, Lawal M, Azeez RA, Afolabi AZ. Comparative analysis of back-propagation neural network and K-means clustering algorithm in fraud detection in online credit card transactions. Fountain Journal of Natural and Applied Sciences. 2019;8(1):315. doi: 10.53704/fujnas.v8i1.315.

[24]. Meenu et al. (2020) Meenu, Meenu, Gupta S, Patel S, Kumar S, Chauhan G. Anomaly detection in credit card transactions using machine learning. International Journal of Innovative Research in Computer Science & Technology (IJIRCST) 2020;8(3):2020. doi: 10.21276/ijircst.2020.8.3.5

[25]. Vijayakumar et al. (2020) Vijayakumar V, Divya NS, Sarojini P, Sonika K. Isolation forest and local outlier factor for credit card fraud detection system. International Journal of Engineering and Advanced Technology (IJEAT) 2020;9(4):261–265. doi: 10.35940/ijeat.D6815.049420

[26]. Palekar et al. (2020) Palekar V, Kharade S, Zade H, Ali S, Kamble K, Ambatkar S. Credit card fraud detection using isolation forest. International Research Journal of Engineering and Technology (IRJET) 2020;7(3):1–6

[27]. Harwani et al. (2020) Harwani H, Jain J, Jadhav C, Hodavdekar M. Credit card fraud detection technique using hybrid approach: an amalgamation of self organizing maps and neural networks. International Research Journal of Engineering and Technology (IRJET) 2020;7(7):2020.

[28]. Deb, Ghosal & Bose (2021) Deb K, Ghosal S, Bose D. A comparative study on credit card fraud detection. EngrXiv preprint. 2021.

[29]. Dzakiyullah, Pramuntadi & Fauziyyah (2021) Dzakiyullah NR, Pramuntadi A, Fauziyyah AK. Semi-supervised classification on credit card fraud detection using autoencoders. Journal of Applied Data Sciences. 2021;2(1):1–7. doi: 10.47738/jads.v2i1.16.

[30]. Pratap & Vijayaraghavulu (2021) Pratap BG, Vijayaraghavulu P. A hybrid method for credit card fraud detection using machine learning algorithm. International Journal of Computers, Electrical and Advanced Communication Engineering (IJCEACE) 2021;10(19):46–50.

[31]. Misra et al. (2020) Misra S, Thakur S, Ghosh M, Saha SK. An autoencoder based model for detecting fraudulent credit card transaction. Procedia Computer Science. 2020;167:254–262. doi: 10.1016/j.procs.2020.03.219.

[32]. Wu, Cui & Welsch (2020) Wu E, Cui H, Welsch RE. Dual autoencoders generative adversarial network for imbalanced classification problem. IEEE Access. 2020;8:91265–91275. doi: 10.1109/ACCESS.2020.2994327.

[33]. Makolo & Adeboye (2021) Makolo A, Adeboye T. Credit card fraud detection system using machine learning. International Journal of Information Technology and Computer Science. 2021;2021(4):24–37. doi: 10.5815/ijitcs.2021.04.03.

[34]. Daliri (2020) Daliri S. Using harmony search algorithm in neural networks to improve fraud detection in banking system. Computational Intelligence and Neuroscience. 2020;2020(1):1–5. doi: 10.1155/2020/6503459

[35]. Hussein, Abbas & Mahdi (2021) Hussein NK, Abbas AR, Mahdi BS. Fraud classification and detection model using different machine learning algorithm. Tech-Knowledge. 2021;1(1):2021.