Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Financial Fraud Detection Based on Machine and Deep Learning: A Review

Rojan Zaki Abdulkreem[1,2], Adnan Mohsin Abdulazeez[1]
[1]zakiroj@gmail.com, [2]adnan.mohsin@dpu.edu.krd
[1]Duhok Polytechnic University, Kurdistan Region – Iraq
[2]Akre University for Applied Science / Technical College of Informatics – Akre / Department of Information Technology – Kurdistan Region – Iraq

Article Information

Submitted : 23 May 2024
Reviewed : 14 Jun 2024
Accepted : 30 Jun 2024

Keywords

Machine Learning , Deep Learning, Fraud Detection, Neural Networks, Anomaly Detection

Abstract

Financial fraud detection is crucial for protecting the integrity of financial markets and institutions globally. Recent advancements in machine learning (ML) and deep learning (DL) have dramatically enhanced the ability to detect and prevent fraudulent activities across various sectors. This review paper examines the implementation of ML and DL in fraud detection, highlighting the evolution from traditional methods to sophisticated models like neural networks, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). We explore different ML techniques such as supervised, unsupervised, and hybrid approaches, their effectiveness in handling large, imbalanced datasets, and their application in real-world scenarios. Special attention is given to the integration of technologies like blockchain and IoT with AI to innovate fraud detection frameworks. Despite the promising advancements, challenges remain, such as the need for large volumes of labeled data, potential model bias, and the black-box nature of many deep learning models. Future directions focus on enhancing model transparency, addressing privacy concerns, and expanding the use of federated learning. This review aims to demonstrate the effectiveness of current technologies and encourage their adoption in enhancing global financial security.

A. Introduction

Financial fraud detection has become a critical area of focus in recent years, with the rise of sophisticated fraudulent activities in various financial sectors. Machine learning and deep learning techniques have emerged as powerful tools in combating financial fraud by enabling the development of advanced detection models. Researchers have explored a range of innovative approaches to enhance fraud detection accuracy and efficiency.

Researchers have employed deep learning algorithms like recurrent neural networks (RNN) and long short-term memory (LSTM) to create models for detecting financial statement fraud [1][2]. Furthermore, machine learning techniques like support vector machines and decision trees have been utilized to pinpoint anti-money laundering activities within the healthcare sector [3][4][5]. Additionally, the use of oversampling techniques like Synthetic Minority Over-sampling Technique (SMOTE) has been proposed to address data imbalance issues in credit card fraud detection models, leading to improved accuracy compared to traditional methods [6].

The integration of artificial intelligence (AI) and deep learning has played a pivotal role in identifying fraudulent financial transactions, money laundering schemes, and bank transfer scams [7]. Studies have also honed in on credit card fraud detection by leveraging deep learning models based on autoencoders and convolutional neural networks [8][9]. Through a comparison of machine learning and deep learning techniques, researchers have strived to identify the most effective algorithms for spotting fraudulent transactions within real-world credit card datasets [10][11].

Moreover, researchers have proposed the development of deep convolutional neural network models for enhancing the accuracy of credit card fraud detection and efficiently managing large volumes of data [12][13]. Investigations have also explored the utilization of cutting-edge deep learning models to detect financial fraud in Chinese listed companies by amalgamating numerical features from financial statements and textual data [14]. Additionally, data mining techniques have proven instrumental in detecting various types of financial fraud, employing methods such as logistic regression, decision trees, support vector machines, neural networks, and naïve Bayes [15].

The collaboration between academia and industry has been pivotal in driving innovations in financial fraud detection [16][17]. Researchers continuously experiment with hybrid models that combine multiple machine learning techniques to leverage their unique strengths, enhancing the overall detection capabilities [18]. For instance, combining decision trees with neural networks can exploit the interpretability of decision trees and the predictive power of neural networks [19][20]. This interdisciplinary approach not only refines the accuracy of detection models but also caters to the specific requirements of different financial sectors, ranging from banking to insurance [21].

This continuous innovation in financial fraud detection is also complemented by efforts to improve data quality and accessibility. With the increase in digital financial transactions, ensuring the integrity and security of data has become paramount [22][23]. Researchers are focusing on developing protocols and algorithms that can effectively preprocess, clean, and secure data before it's used

in fraud detection models. This not only enhances the performance of the detection systems but also helps in maintaining compliance with global data protection regulations. Such advancements in data management are crucial for enabling more accurate and timely detection of financial fraud, further bolstering the reliability and efficiency of these systems [24][25].

This review aims to focus on enhancing global adoption of advanced fraud detection technologies. By showcasing the effectiveness of these systems, the goal is to encourage their widespread use, especially in under-equipped regions. Training is also prioritized to ensure personnel can effectively manage and trust these tools. This initiative aims to strengthen financial systems worldwide against evolving fraud tactics, thereby supporting global economic stability and consumer confidence.

The remainder of this paper is structured as follows: Section 2 delves into the results of the review analysis, meticulously detailing the findings and underscoring their significance in the context of existing research. Section 3 presents a synthesis of the most crucial conclusions drawn from the analysis, offering insights into the implications and potential impact of the study. In Section 4, a comprehensive discussion of the review analysis is provided, elaborating on the nuances and broader ramifications of the results. Finally, Section 5 concludes the paper by summarizing the research findings, encapsulating the essence of the study and its contributions to the field.

B. Research Method

2.1 The Scope and Impact of Financial Fraud

Financial fraud poses a significant threat to the integrity and stability of financial markets and institutions around the world. It encompasses a wide range of illegal activities, including credit card fraud, insurance fraud, securities fraud, and banking fraud, each capable of causing substantial financial losses and eroding trust in financial systems. As technology advances and financial systems become more integrated globally, fraud tactics continually evolve, presenting ongoing challenges that require increasingly sophisticated detection and prevention strategies.

2.2 Machine Learning in Fraud Detection

The advent of machine learning has introduced more dynamic and adaptable approaches to fraud detection. Unlike traditional methods, machine learning algorithms can learn and evolve in response to new data, improving their accuracy and effectiveness over time. Supervised learning techniques, such as decision trees, support vector machines, and ensemble methods like random forests, have been widely adopted for fraud detection, benefiting from their ability to handle large, complex datasets and their effectiveness in classifying imbalanced data [26].

Unsupervised learning techniques, including k-means clustering and principal component analysis, are used to detect anomalies in data without prior labeling. These methods are particularly useful in identifying unknown types of fraud. Hybrid approaches that combine multiple machine learning models aim to leverage the strengths of various algorithms to improve detection rates and reduce false positives [27][28].

2.3 The Role of Deep Learning in Fraud Detection

Deep learning, a powerful subset of machine learning, has become increasingly significant in fraud detection due to its ability to process and learn from large, complex datasets. Particularly, neural network architectures like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are highly effective in detecting subtle and complex patterns indicative of fraud. CNNs, as demonstrated in Figure 1, excel in automatically extracting features from vast and varied data without human input, giving them a distinct advantage in dynamic environments like financial fraud detection. This capability allows deep learning models to surpass traditional machine learning models in many aspects, adapting efficiently to the ever-evolving tactics of financial fraud [29][30].

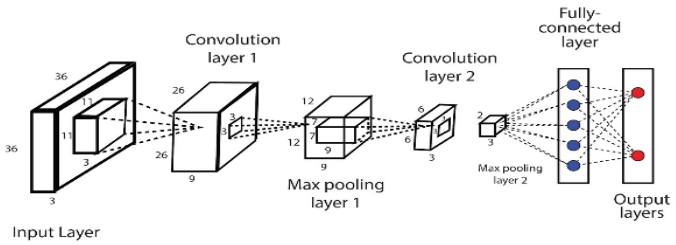

Figure 1. CNN architecture showing layers from input to output, used in fraud detection.

2.4 Advances in Machine and Deep Learning for Fraud Detection

Recent advancements in machine learning and deep learning technologies have significantly bolstered fraud detection capabilities. Innovations such as real-time processing of large datasets, advanced anomaly detection algorithms that adapt to new and emerging fraud tactics, and the integration of AI with other technologies like blockchain and IoT are transforming the landscape of fraud detection. These advances enable more sophisticated, efficient, and scalable fraud detection systems that are capable of learning from the data in real time and adjusting their detection mechanisms dynamically [31][32].

2.5 Advancement's mains in Machine and Deep Learning of Future Directions

Recent advancements in neural network designs have significantly enhanced the capability to detect and analyze fraud through personalized data processing [33]. The architecture shown in Figure 2 features multiple input feature sets that are processed through complex layers—including batch normalization and ReLU activation followed by dropout—to effectively manage overfitting and improve model generalization. This tailored approach allows for outputs that are specific to individual user behaviors, making it a potent tool in the detection of sophisticated fraud schemes that may vary from user to user [34].

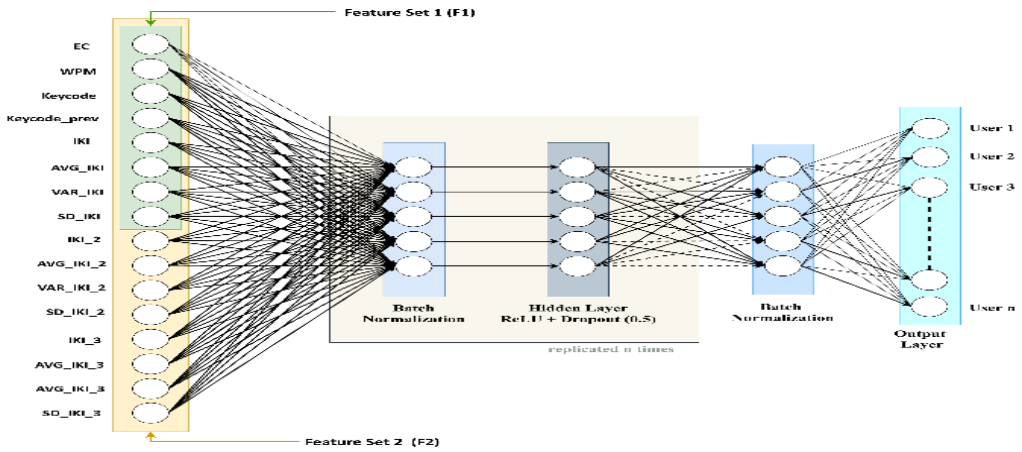

Figure 2. Neural network architecture with multi-feature input and user-specific output layers, illustrating personalized fraud detection capabilities

2.6 Challenges and Future Directions

Despite their advantages, both machine learning and deep learning face significant challenges in fraud detection. These include dealing with imbalanced datasets where instances of fraud are rare, the need for large volumes of labeled data for training, and the potential for model bias. Furthermore, the black-box nature of many deep learning models poses challenges for transparency and explain ability, critical factors in regulatory and compliance contexts.

Emerging trends, such as explainable AI, adversarial machine learning, and federated learning, offer promising avenues for future research. These technologies aim to enhance the trustworthiness, robustness, and effectiveness of fraud detection systems while addressing privacy concerns and adapting to new types of fraud.

C. Related Works

Ding et al. in 2023 introduced an improved Variational Autoencoder Generative Adversarial Network (VAEGAN) to enhance credit card fraud detection, focusing on synthetic data generation to address class imbalances in training. By comparing several models and algorithms, VAEGAN demonstrated superior performance in precision and F1_score, highlighting its effectiveness in managing skewed datasets typical in fraud detection tasks. However, the complexity of the improved VAEGAN model may lead to overfitting and computational inefficiencies. To mitigate this, simplifying the model by reducing layers or encoders and implementing regularization techniques could improve efficiency and prevent overfitting [35].

Ishida et al. introduced SA-PatchCore, an anomaly detection model integrating self-attention with the PatchCore model to enhance detection of co-occurrence relationship anomalies. The model, tested using a specially created Co-occurrence Anomaly Detection Screw Dataset (CAD-SD), shows high performance in identifying both local and co-occurrence anomalies compared to standard

PatchCore, particularly excelling in complex detection scenarios where traditional models falter. The complexity of the SA-PatchCore model may lead to higher computational demands and overfitting in simpler anomaly detection tasks. Optimize the model by adjusting the self-attention mechanism to reduce computational demands without sacrificing detection accuracy [36].

Teuku Rizky Noviandy et al. (2023) explored the use of the XGBoost algorithm alongside SMOTE-ENN data augmentation techniques to address credit card fraud detection in imbalanced datasets. This methodology significantly improved the balance between precision and recall, proving to be vital for enhancing contemporary financial management systems. The study underscores the potential to strengthen financial integrity and boost consumer trust through more accurate fraud detection. However, the research's focus on a specific algorithm and augmentation technique limits its generalizability. To overcome this limitation, expanding the range of tested algorithms and data augmentation methods is suggested, potentially enhancing the robustness and wider applicability of the results [37].

Debener et al. (2023) explored the effectiveness of unsupervised and supervised machine learning methods—specifically isolation forests and XGBoost—for detecting insurance fraud, using data from a German insurance company. Both methods were found to be effective, each identifying unique cases of fraud, highlighting their potential to enhance fraud detection capabilities within the insurance industry. However, the study's reliance on a single dataset from one type of insurance and limited machine learning methods may affect its generalizability. To overcome this limitation, expanding the research to include diverse datasets from various types of insurance and incorporating additional machine learning methods could improve the robustness and applicability of the findings [38].

Yoo et al. (2023) compared Medicare fraud detection methods using machine learning and Graph Neural Networks (GNNs), focusing on graph centrality measures within provider-beneficiary networks. They demonstrated that traditional machine learning models augmented with graph centrality features significantly outperformed GNNs, showing improvements in precision, recall, and F1-score. This enhancement suggests substantial potential cost savings in fraud prevention. However, the study's reliance on specific graph centrality measures and data from a single region may limit its generalizability. To address this limitation, it is suggested to broaden the dataset diversity and explore additional centrality measures to strengthen the model's robustness and wider application [39].

Labu and Ahammed (2024) investigated the integration of AI and ML into cyber threat detection, demonstrating that Random Forest algorithms, with an 83.94% accuracy rate, outperformed other models in identifying cyber threats. Their study emphasizes the effectiveness of these technologies in enhancing real-time fraud detection and accurate transaction identification within financial institutions. However, the research may have limited generalizability due to its focus on specific technologies and datasets. To address this, expanding the range of technologies and datasets could improve the study's broader applicability and relevance in diverse cybersecurity contexts [40].

Aljabri and Mohammad (2023) investigated click fraud in online advertising by applying various machine learning models to distinguish between human and automated bot interactions. Their research utilized a dataset of user web behaviors, evaluating algorithms like Decision Tree, Support Vector Machine, Naive Bayes, and Random Forest. The Random Forest algorithm emerged as the most effective, achieving the highest accuracy in all evaluated metrics. However, the study faced limitations due to potential bot activities within the dataset, which might compromise the accuracy in representing genuine human behaviors. To overcome this, they suggested future research should implement more stringent verification techniques to ensure the exclusion of bots, thereby enhancing dataset reliability and model accuracy [41].

Hassan Najadat et al. (2020) developed a credit card fraud detection system using machine and deep learning models on the IEEE-CIS Fraud Detection dataset from Kaggle. The study applied several classifiers like Naive Bayes, Random Forest, Decision Trees, and advanced deep learning models including BiLSTM and BiGRU. The hybrid model combining BiLSTM and BiGRU achieved the highest accuracy, demonstrating superior performance at 91.37%. This approach highlights the effectiveness of integrating machine learning with deep learning techniques for fraud detection. However, the study's limitation lies in its reliance on a single dataset, which may not capture the full diversity of real-world transactions. Expanding the dataset used could enhance the model's effectiveness and general applicability [42].

Schneider and Brühl (2023) investigated accounting fraud detection in publicly listed U.S. firms, utilizing machine learning to analyze the influence of CEO characteristics alongside raw financial data. Employing algorithms like Random Forest and XGBoost, their research highlighted that nonlinear model are particularly effective, capturing complex relationships between CEO traits and fraud occurrences. Key findings suggest that CEO Network Size and CEO Age significantly impact fraud prediction accuracy, emphasizing the enhanced performance of models integrating CEO and financial data over those using financial data alone. The study's focus on U.S. publicly listed companies, however, might limit its generalizability to other contexts or regions, suggesting the need for broader data inclusion to improve model robustness and applicability [43].

Agarwal (2023) developed a machine learning method for fraud detection in medical claim insurance, focusing on the K-means clustering algorithm. This unsupervised approach effectively groups similar data points to detect patterns indicative of fraud, resulting in significant enhancements in detection metrics such as accuracy, precision, recall, and F1-score compared to traditional methods. However, the unsupervised nature of the model may misclassify legitimate atypical claims as fraudulent, leading to potential false positives. Integrating semi-supervised learning could mitigate this issue by utilizing both labeled and unlabeled data, improving the model's accuracy and reducing false positives [44].

Sina Ahmadi (2023) explores the use of OpenAI's technologies in fraud detection within the financial sector, highlighting the rising complexity of fraud schemes that necessitate sophisticated solutions like machine learning algorithms, including decision trees, logistic regression, and neural networks. These technologies have proven effective in enhancing the accuracy and efficiency of

fraud detection systems. However, the study also acknowledges the dual-use nature of these AI tools, as they can be exploited by fraudsters to conduct sophisticated scams. This underscores the need for a balanced approach to utilizing AI in combating financial crimes. The paper suggests that while OpenAI's tools are beneficial, a broader range of AI technologies should be considered to fully address the diverse challenges in fraud detection, thereby enhancing the effectiveness of these systems [45].

Abdu Salam and colleagues (2023) advanced the security of smart manufacturing by integrating anomaly detection with Zero-Knowledge Proofs (ZKPs), specifically using zk-SNARKs. Their approach utilized deep learning architectures for anomaly detection in smart manufacturing systems, achieving high detection accuracy for conditions like temperature and pressure irregularities. Verification of these anomalies through ZKPs ensured data confidentiality while maintaining high integrity, demonstrating an impressive success rate. However, the practical application of such sophisticated cryptographic methods may be limited by their computational demands. A potential improvement would be to streamline ZKP algorithms or to use hardware acceleration, which could enhance the feasibility of this approach for wider implementation in industry settings [46].

Sandeep Dasari and Rajesh Kaluri (2023) investigated the classification of DDoS attacks using hierarchical machine learning models enhanced by hyperparameter optimization on the CICIDS 2017 dataset. Their approach incorporated algorithms such as XGBoost, LGBM, CatBoost, Random Forest, and Decision Tree, with preprocessing that included min-max scaling and SMOTE for data balancing. LASSO was employed for feature selection, which pinpointed key attributes for model training. The study highlighted the LGBM classifier's superior performance, achieving an impressive 99.77% accuracy, showcasing its potential in detecting DDoS attacks effectively. However, the reliance on a single dataset may affect the generalizability of these results. Expanding the study to incorporate multiple datasets from varied network environments is suggested to improve the findings' robustness and application [47].

In 2021, Khaled Gubran Al-Hashedi and Pritheega Magalingam conducted a comprehensive review of financial fraud detection using data mining techniques from 2009 to 2019, examining the application of these methods across different types of fraud such as banking, insurance, and cryptocurrency. The study emphasizes the effectiveness of SVM, Naïve Bayes, and Random Forest in combating fraud and outlines the predominant use of these techniques in banking and insurance sectors. This review contributes valuable insights into the evolution and efficacy of data mining in financial fraud detection over a decade. However, it primarily synthesizes existing research without new empirical data, which may limit the practical application of the findings. To enhance the relevance and applicability of the review, it is suggested to conduct empirical studies using these techniques on contemporary, real-world datasets to validate their effectiveness in current fraud detection scenarios [48].

Fatima Rashed Alzaabi and Abid Mehmood (2023) provide a comprehensive review of machine learning methods for detecting malicious insider threats, highlighting the superiority of deep learning and natural language processing

techniques over traditional methods. Their analysis, primarily using the CMU CERT dataset, showcases these advanced methods' effectiveness in identifying subtle and complex insider behaviors. The paper also suggests that incorporating time-series-based techniques could further enhance detection capabilities. However, the study's focus on a specific dataset may restrict the generalizability of its findings. To address this limitation, they recommend expanding research to include a broader range of real-world scenarios and datasets, which could improve the robustness and applicability of insider threat detection strategies [49].

Faisal S. Alsubaei and colleagues (2023) developed a hybrid deep learning framework for phishing detection, leveraging a ResNeXt-embedded Gated Recurrent Unit (RNT) model optimized with the Jaya method. By integrating SMOTE for data balancing and employing advanced feature extraction via autoencoders and ResNet, the framework achieves significant improvements in detection efficiency and accuracy, surpassing traditional methods by 11% to 19%. Tested on real phishing datasets, the model demonstrated high accuracy and low false positive/negative rates. However, the reliance on complex deep learning techniques and substantial computational resources could limit its broader applicability. To address this, adopting more computationally efficient models could expand the framework's usability across different computational platforms [50].

B. Dangsawang and S. Nuchitprasitchai (2024) introduced the SHIELD model for detecting customs fraud using unstructured social media data, employing Logistic Regression, Gated Recurrent Unit (GRU), and Long Short-Term Memory (LSTM) algorithms. This model effectively categorizes commercial goods into suspicious and non-suspicious categories by analyzing data from Twitter and Facebook. Among the techniques tested, LSTM displayed the highest accuracy and F1-score, indicating its superior capability in identifying potential fraud. The reliance on social media data poses a challenge due to variability in data quality and completeness. To address this, the study suggests enhancing data reliability and expanding verification processes to improve the model’s effectiveness and applicability in real-world scenarios [51].

Abdul Wahid and colleagues (2023) introduced a real-time telecom fraud detection system using a Neural Factorization Autoencoder (NFA), incorporating Neural Factorization Machines (NFM) and an Autoencoder (AE) to analyze customer calling patterns. The model includes a memory module that adapts to changing customer behaviors, offering a dynamic approach to fraud detection. Tested on a substantial real-world dataset, the NFA demonstrated superior performance, achieving an AUC of 91.06% and an F1-score of 95.45%, outperforming existing methods. However, its reliance on complex neural architectures may hinder scalability and increase operational costs. To address this, future enhancements could focus on optimizing the model to reduce computational demands while maintaining high detection accuracy [52].

Zainab Saad Rubaidi and colleagues (2023) focused on vehicle insurance fraud detection, employing various supervised machine learning algorithms and data resampling techniques, such as NearMiss, SMOTE, and a hybrid approach involving multiple oversampling methods. Their comparative study highlighted the Random Forest model using the hybrid data augmentation approach, which

achieved an F1-score and accuracy of 0.975, indicating high efficacy in detecting fraudulent activities. The research showed that integrating diverse oversampling strategies could enhance model performance in fraud detection tasks. However, the potential over-representation of minority class features due to the hybrid augmentation approach might lead to overfitting. To address this, incorporating more robust validation and regularization methods could help improve the model's accuracy and generalizability across various datasets [53].

Danial Jamil and colleagues (2023) developed machine learning models to enhance fraud detection in green finance, focusing on the challenges of imbalanced data using the PaySim dataset. They applied various algorithms, including Random Forests, Recurrent Neural Networks, and K-Nearest Neighbors, and explored deep learning techniques like Long Short-Term Memory (LSTM) and Artificial Neural Networks (ANN) to reveal hidden patterns in transactions. Their research achieved enhanced precision and effectiveness in fraud detection, significantly mitigating fraud-related losses and maintaining consumer trust. However, the use of a synthetic dataset like PaySim might not fully replicate real-world transaction dynamics. To enhance the models' applicability and realism, incorporating real transaction data from various financial settings is recommended [54].

Fatima Adel Nama and colleagues (2024) utilized Recurrent Neural Networks (RNNs) to enhance fraud detection in mobile money transactions, employing a synthetic dataset generated by PaySim. Their model effectively captured the temporal dynamics of transaction data to identify fraudulent patterns, demonstrating its robust capabilities with an accuracy of 99.87% and an F1-score of 0.99. This high performance indicates its potential utility in distinguishing between legitimate and fraudulent transactions efficiently. However, the reliance on synthetic data might limit the model's effectiveness in real-world scenarios. To address this, incorporating real transactional data could improve both the accuracy and practical applicability of the model, ensuring it remains effective against evolving financial fraud techniques [55].

Ali Raza and colleagues (2023) introduced a novel machine learning methodology, the Class Probability Random Forest (CPRF), to enhance network attack detection. Utilizing the CICIDS2017 dataset, this approach involves generating class probability features to improve the effectiveness of detection models. They demonstrated that the CPRF method, combined with Random Forest, significantly outperformed conventional techniques, achieving an accuracy of 99.9%. Validated through k-fold cross-validation and optimized with hyperparameter tuning, this method showcases substantial improvements over traditional network intrusion detection approaches. However, the reliance on a single dataset might restrict the general applicability of the findings. To overcome this, incorporating more diverse datasets and attack scenarios is suggested to ensure the method's robustness and applicability in various settings [56].

In 2023, Amirul Islam and colleagues developed the Credit Card Anomaly Detection (CCAD) model, an advanced ensemble-based machine learning approach for detecting anomalies in credit card transactions. Utilizing a combination of four outlier detection algorithms as base learners and an XGBoost algorithm as the meta-learner, the model addresses the challenges of imbalanced and overlapping class samples effectively. Tested using stratified sampling and k-fold cross-

validation on credit card datasets, the CCAD model demonstrated superior performance, significantly outperforming existing models in detecting anomalies, especially from minority classes. However, the model's reliance on complex algorithms and multiple layers increases its computational demands. Optimizing the ensemble components and employing more efficient algorithms could potentially reduce its complexity while maintaining high detection accuracy [57].

In 2024, R. Jayaraj and team developed a new phishing detection system using a machine learning approach that leverages a Hybrid Ensemble Feature Selection (HEFS) method. This method employs a novel Cumulative Distribution Function gradient (CDF-g) algorithm alongside data perturbation techniques to optimize feature selection, significantly enhancing the identification of phishing URLs. The system was tested and demonstrated high accuracy in distinguishing between legitimate and malicious websites, showcasing its efficacy in real-time phishing detection. Despite its strengths, the system's reliance on synthetic or preset datasets might limit its ability to handle real-world phishing scenarios. To improve the system's practical application, incorporating real-time data from varied sources is recommended, which could adapt more effectively to evolving phishing strategies [58].

In 2023, Mei-See Cheong, Mei-Chen Wu, and Szu-Hao Huang introduced an advanced stock anomaly detection system utilizing Spatio-Temporal Convolutional Neural Networks combined with a Relation Network (STCNN-RN), further optimized by a Genetic Algorithm. This innovative system analyses multiple financial time-series data to effectively detect anomalies in stock market behaviors. Through extensive testing across various financial markets, the model demonstrated high accuracy in identifying market irregularities, offering valuable insights for investors. Despite its effectiveness, the complex nature of the model demands significant computational resources, which may hinder practical deployment. To enhance usability and reduce computational costs, it is suggested to explore optimization strategies and develop simpler model variants for broader application in real-time financial environments [59].

In 2023, Chandana Gouri Tekkali and Karthika Natarajan introduced the RDQN model, an innovative approach that integrates deep reinforcement learning with rough set theory for enhancing digital transaction fraud detection. The model employs rough set theory for effective feature selection and utilizes a Deep Q-Network (DQN) for the classification of fraudulent activities, resulting in notable improvements in both accuracy and processing speed. This method was particularly developed to address the challenges of detecting increasingly sophisticated fraudulent transactions. However, the model's effectiveness is currently limited by its reliance on a specific dataset that may not capture global transaction variability. To extend its utility and ensure more comprehensive fraud detection, it is essential to test the RDQN model using a broader variety of transaction data from multiple geographic regions and different platforms [60].

In 2021, Benchaji et al. introduced a sophisticated credit card fraud detection system using LSTM networks combined with an attention mechanism, designed to enhance the detection accuracy by leveraging the sequential nature of transaction data. The system incorporates dimensionality reduction techniques like UMAP and SMOTE for addressing imbalanced datasets, enabling the model to focus on the

most relevant features for predicting fraud. Tested across multiple datasets, this model demonstrated significant improvements over traditional methods, showcasing its potential in the practical application of fraud detection in the financial sector. However, the dependency on LSTM may limit handling of complex and long sequence dependencies. Incorporating advanced architectures like Transformers could offer improvements in handling these complexities, enhancing both performance and scalability [61].

Rejwan Bin Sulaiman and colleagues (2023) explored credit card fraud detection by integrating Artificial Neural Networks (ANN) within a federated learning framework to form a hybrid machine learning system. This approach aimed to enhance detection accuracy while preserving data privacy, using a combination of Random Forest, Support Vector Machines, and federated learning to tackle the challenges associated with maintaining data confidentiality in fraud analytics. The hybrid model showed improved performance on synthetic datasets by demonstrating robust detection capabilities. However, its application might be limited by the use of such controlled datasets, which may not capture the full spectrum of real-world fraudulent activities. To bolster the model's practicality and adaptability, incorporating real-world data from diverse financial systems worldwide could significantly enhance its effectiveness and reliability [62].

In 2023, Hashemi, Mirtaheri, and Greco enhanced credit card fraud detection by employing Bayesian optimization to fine-tune hyperparameters across several machine learning models on unbalanced data. They utilized LightGBM, XGBoost, and CatBoost, and implemented majority voting ensemble learning alongside deep learning improvements. Their methodology significantly enhanced detection capabilities, demonstrated by achieving ROC-AUC of 0.95, precision of 0.80, and an F1 score of 0.81. This was accomplished through extensive experiments on real-world datasets, particularly highlighting the effectiveness of their ensemble approach. However, the reliance on traditional models and static public datasets may limit the system's responsiveness to new and evolving fraud techniques. To counter this, integrating real-time transaction data and adaptive learning models could potentially improve the robustness and adaptiveness of their fraud detection system [63].

Suraya Nurain Kalid and colleagues (2024) conducted a systematic review of machine learning techniques for detecting credit card fraud and payment defaults, specifically addressing the challenges associated with imbalanced class distribution and class overlap. They highlighted the potential of deep learning, ensemble learning, and sampling methods to effectively manage these issues. The study also recommended using performance metrics like True Positive Rate (TPR) and Area Under the Curve (AUC) to evaluate these techniques. The review demonstrated that these advanced methods significantly enhance fraud detection capabilities. However, the study’s limitation lies in its lack of direct empirical testing, which could impact its practical implementation. To overcome this, conducting empirical research on diverse datasets is suggested to confirm the effectiveness of these techniques in real-world settings [64].

In 2023, Taha and Malebary developed an optimized light gradient boosting machine (OLightGBM) enhanced by Bayesian-based hyperparameter optimization for credit card fraud detection. Tested across two real-world datasets, the

OLightGBM model demonstrated superior performance, achieving an accuracy of 98.40%, AUC of 92.88%, precision of 97.34%, and an F1-score of 56.95%, thus significantly outperforming traditional fraud detection methods. This study illustrates the potential of advanced machine learning techniques in addressing the complex challenges of fraud detection in the financial sector. However, the model's effectiveness might be limited by its reliance on specific datasets, which may not capture the full spectrum of global credit card fraud. To further improve its applicability, expanding the variety of testing datasets could help in generalizing the model's effectiveness across different fraud scenarios and transaction types [65].

In 2024, Zengyi Huang and colleagues explored the application of K-means clustering for financial fraud detection, showcasing its capacity to adaptively identify anomalous patterns in vast amounts of transaction data, significantly outstripping traditional rule-based detection methods. Their research highlighted the method's flexibility and precision, particularly in allocating resources efficiently within financial institutions to focus on high-risk areas, thereby enhancing the overall security and reliability of financial systems. However, the study's main limitation lies in its theoretical approach without real-world empirical testing, which raises questions about its practical effectiveness. To solidify the findings and ensure broader applicability, it is suggested that future studies implement this K-means clustering technique in real operational settings and perform empirical evaluations to verify its efficacy and optimize its deployment in detecting financial fraud [66].

In 2024, Farhan Aslam published a comprehensive review on the advancements of machine learning algorithms in credit card fraud detection, with a particular emphasis on the effectiveness of the Light Gradient Boosting Machine (LGBM). The study provides a detailed comparison between LGBM and traditional machine learning techniques, highlighting LGBM’s superior ability to process large datasets and its proficiency in identifying complex fraud patterns swiftly and accurately. This review points out that while traditional methods falter in handling modern fraud dynamics, LGBM's fast processing times and high predictive accuracy position it as a highly effective tool for financial institutions aiming to combat fraud. However, Aslam’s analysis primarily remains theoretical and lacks empirical validation. To overcome this limitation, it is crucial to conduct practical tests of the LGBM algorithm in real-world fraud detection scenarios to firmly establish its utility and effectiveness[67].

In their 2020 study, Badr Omair and Ahmad Alturki systematically review fraud detection metrics in business processes, focusing on process-based fraud (PBF). They analyze various metrics applied to assess fraud risks in business processes, identifying significant gaps in current methodologies, particularly the lack of comprehensive metrics that address all conceptual aspects of business processes. The paper emphasizes the theoretical foundations of these metrics but highlights a critical shortcoming: the absence of practical applications or empirical validations. Consequently, Omair and Alturki suggest that future research should involve the practical implementation and testing of these metrics in real business scenarios to confirm their efficacy and adaptability in detecting and mitigating

business process fraud effectively. This approach could help bridge the gap between theoretical constructs and real-world applicability in fraud detection [68].

In 2020, Patricia Craja, Alisa Kim, and Stefan Lessmann developed a deep learning model aimed at detecting financial statement fraud. Their approach integrates financial ratios with textual analysis from the Management Discussion and Analysis (MD&A) sections of corporate annual reports. Utilizing a hierarchical attention network (HAN), the model efficiently extracts and prioritizes key textual features, significantly enhancing fraud detection capabilities. This method not only identifies potential fraud but also provides interpretable results by highlighting specific "red-flag" sentences, aiding stakeholders in decision-making processes. However, the model's reliance solely on MD&A text could restrict its effectiveness, as these sections may not consistently contain fraud indicators. To broaden its applicability and improve detection accuracy, the model could be refined to include additional textual segments from corporate reports and other varied data sources [69].

Table 1. Overview of Recent Studies on Fraud Detection Using Machine Learning and Deep Learning Techniques

Ref	Authors	Year	Dataset	Based Model	Technique	Description	Advantages	Limitations
35	Ding et al.	2023	Credit card data	VAEGAN	Synthetic data generation	Enhanced credit card fraud detection with improved VAEGAN model	Superior precision and F1 score; addresses class imbalances	Complexity may lead to overfitting and computational inefficiencies
36	Ishida et al.	2023	CAD-SD	SA-PatchCore	Self-attention	Integration of self-attention with PatchCore for anomaly detection	High performance in complex detection scenarios	Higher computational demands; potential overfitting
37	Teuku Rizky Noviandy et al.	2023	Credit card data	XGBoost	SMOTE-ENN	Use of XGBoost and data augmentation to improve fraud detection	Improved balance between precision and recall	Limited generalizability due to specific techniques
38	Debener et al.	2023	German insurance data	Various	Isolation forests, XGBoost	Explored machine learning methods for insurance fraud	Effective in identifying unique cases of fraud	Reliance on a single dataset and limited methods
39	Yoo et al.	2023	Medicare data	GNNs	Graph centrality	Compared Medicare fraud detection methods focusing on provider-beneficiary networks	Traditional models with graph feature outperformed GNNs	Reliance on specific measures and data from a single region
40	Labu and Ahammed	2024	Cyber Threat Data	Random Forest	Machine learning	Investigated AI and ML in cyber threat detection with a focus on Random Forest.	High accuracy rate in detecting cyber threats.	Limited generalizability due to specific focus.
41	Aljabri and Mohammad	2023	Online Advertising Data	Random Forest	Machine learning	Studied click fraud detection using various ML models to distinguish human/bot interactions.	Random Forest achieved highest accuracy.	Potential bot activities may affect data accuracy.
42	Hassan Najadat et al.	2020	IEEE-CIS Fraud Detection Dataset	BiLSTM, BiGRU	Deep learning, machine learning	Developed a hybrid model combining BiLSTM and	High accuracy of 91.37%.	Reliance on a single dataset.

						BiGRU for credit card fraud detection.		
43	Schneider and Brühl	2023	U.S. Financial Data	Random Forest, XGBoost	Machine learning	Analyzed influence of CEO characteristics on accounting fraud detection.	Effective in capturing complex relationships.	Limited to U.S. publicly listed companies.
44	Agarwal	2023	Medical Claim Data	K-means	Clustering, unsupervised learning	Developed a K-means based method for detecting fraud in medical insurance claims.	Significant enhancements in detection metrics.	May misclassify legitimate atypical claims.
45	Sina Ahmadi	2023	Financial Sector Data	Various	Decision trees, logistic regression	Explores OpenAI's technologies in financial fraud detection.	Enhanced accuracy and efficiency of fraud detection systems.	Dual-use nature of AI tools.
46	Abdu Salam et al.	2023	Smart manufacturing data	Deep learning	Zero-Knowledge Proofs	Integrated anomaly detection with ZKPs in smart manufacturing	High detection accuracy; ensured data confidentiality	High computational demands limit practical application
47	Sandeep Dasari et al.	2023	CICIDS 2017 Data	LGBM, CatBoost	Machine learning, feature selection	Investigated classification of DDoS attacks using hierarchical ML models.	LGBM showed superior performance.	Reliance on a single dataset.
48	Khaled Gubran Al-Hashedi et al.	2021	Credit Card Data	Various	Data mining	Reviewed financial fraud detection using data mining techniques from 2009-2019	Highlighted effectiveness of SVM, Naïve Bayes, Random Forest	Lacks new empirical data
49	Fatima Rashed Alzaabi et al.	2023	CMU CERT dataset	Deep learning	Natural language processing	Reviewed machine learning methods for detecting malicious insider threats	Superiority of deep learning techniques in identifying complex behaviors	Focus on a specific dataset restricts generalizability
50	Faisal S. Alsubaei et al.	2023	Phishing Data	ResNeXt-embedded GRU	Deep learning, data balancing	Developed a hybrid deep learning framework for phishing detection.	High accuracy and low false rates.	Complex deep learning techniques, high computational cost.
51	B. Dangsawang et al.	2024	Social media data	Various	Logistic Regression, GRU, LSTM	Detected customs fraud using social media data	High accuracy and F1-score with LSTM	Challenge due to variability in social media data quality
52	Abdul Wahid et al.	2023	Telecom data	NFA	Neural Factorization	Introduced a real-time fraud detection system using a novel autoencoder	Dynamic approach with high performance	Complexity may hinder scalability and increase costs
53	Zainab Saad Rubaidi et al.	2023	Vehicle insurance data	Random Forest	Data resampling	Used various machine learning algorithms and resampling techniques	High efficacy in detecting fraudulent activities	Potential overfitting due to over-representation of minority classes
54	Danial Jamil et al.	2023	PaySim Dataset	Various, including LSTM	Deep learning, machine learning	Enhanced fraud detection in green finance using multiple algorithms and deep learning.	Mitigated fraud-related losses effectively.	Use of synthetic dataset may not reflect real transaction dynamics
55	Fatima Adel Nama et al.	2024	PaySim dataset	RNNs	Deep learning	Utilized RNNs for fraud detection in mobile money	High accuracy and robust capabilities	Reliance on synthetic data limits real-world effectiveness

						transactions		
56	Ali Raza et al.	2023	CICIDS2017 dataset	CPRF	Class probability features	Introduced CPRF to enhance network attack detection	Significant improvement over conventional techniques	Reliance on a single dataset may limit general applicability
57	Amirul Islam et al.	2023	Credit card data	CCAD	Ensemble learning	Developed an ensemble-based approach for anomaly detection in credit card transactions	Superior performance in detecting anomalies from minority classes	High computational demands due to complexity
58	R. Jayaraj et al.	2024	Synthetic or preset datasets	HEFS	Feature selection	Developed a phishing detection system using a novel feature selection method	High accuracy in real-time phishing detection	Reliance on synthetic datasets limits handling of real-world scenarios
59	Mei-See Cheong et al.	2023	Financial time-series data	STCNN-RN	Genetic Algorithm	Introduced an advanced stock anomaly detection system	High accuracy in identifying market irregularities	Significant computational resources required
60	Chandana Gouri Tekkali et al.	2023	Digital transaction data	RDQN	Rough set theory, DQN	Enhanced digital transaction fraud detection with deep reinforcement learning	Notable improvements in accuracy and processing speed	Effectiveness limited by reliance on specific dataset
61	Benchaji et al.	2021	Multiple credit card datasets	LSTM	Attention mechanism	Enhanced credit card fraud detection with LSTM and attention mechanism	Significant improvements in detection accuracy	May struggle with complex and long sequence dependencies
62	Rejwan Bin Sulaiman et al.	2023	Synthetic datasets	ANN	Federated learning	Integrated ANN within a federated learning framework for fraud detection	Enhanced detection accuracy while preserving data privacy	Limited by the use of controlled datasets
63	Hashemi, Mirtaheri, and Greco	2023	Real-world datasets	Various	Ensemble learning	Employed Bayesian optimization for fraud detection in credit cards	Enhanced detection capabilities with ROC-AUC of 0.95	Reliance on static datasets may limit responsiveness to new fraud
64	Suraya Nurain Kalid et al.	2024	N/A	Various	Deep learning, ensemble learning	Systematic review of machine learning techniques for fraud detection	Highlighted potential of advanced methods	Lack of direct empirical testing may impact practical implementation
65	Taha and Malebary	2023	Real-world datasets	OLightGBM	Bayesian optimization	Developed an optimized light gradient boosting machine for fraud detection	Superior performance with high accuracy and AUC	Effectiveness might be limited by reliance on specific datasets
66	Zengyi Huang et al.	2024	Two real-world datasets	OLightGBM	Bayesian-based hyperparameter optimization	Developed an optimized light gradient boosting machine for credit card fraud detection.	Superior performance with high accuracy and AUC	Reliance on specific datasets may limit global applicability
67	Farhan Aslam	2024	N/A	Light Gradient Boosting Machine (LGBM)	Machine learning	Reviewed the advancements of LGBM in credit card fraud detection	Fast processing times and high predictive accuracy	Lacks empirical validation
68	Badr Omair and Ahmad Alturki	2020	Business Processes Data	N/A	N/A	Systematically reviewed fraud	Analysis of various	Lack of practical applications or

						detection metrics in business processes	metrics for assessing fraud risks	empirical validations
69	Patricia Craja et al.	2020	Corporate annual reports	HAN	Textual analysis	Developed a deep learning model to detect financial statement fraud	Efficient extraction of key textual features; interpretable results	Reliance solely on MD&A text may restrict effectiveness

D. Discussion

The ML and DL techniques has revolutionized the field of financial fraud detection, providing a more robust, scalable, and efficient means to tackle the complex dynamics of fraudulent activities. This review underscores the significant advancements in algorithmic strategies and model architectures, like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), that have proven to be exceptionally effective in capturing subtle patterns and anomalies within large data sets. These techniques not only enhance the detection capabilities but also offer the flexibility to adapt to new and evolving fraud tactics dynamically.

Hybrid approaches, combining various ML models, capitalize on the strengths of each to address the shortcomings of single-model systems. This versatility is crucial in handling the challenges posed by highly imbalanced datasets prevalent in financial fraud scenarios. Moreover, the adoption of technologies like blockchain and the Internet of Things (IoT) alongside AI has paved the way for creating more secure and transparent systems, mitigating the risks of sophisticated fraud schemes.

However, the journey is not without challenges. The black-box nature of many deep learning models raises concerns about transparency and explain ability, which are critical in regulatory and compliance contexts. Furthermore, the requirement for extensive labeled data sets for training these models poses a significant hurdle, compounded by the potential for model bias and the ethical implications associated with AI decision-making.

Looking ahead, the field is ripe for innovations that enhance the trustworthiness and robustness of fraud detection systems. Emerging trends such as explainable AI, adversarial machine learning, and federated learning promise to address privacy concerns, improve model reliability, and extend the adoption of these technologies globally. The collaborative efforts between academia and industry play a pivotal role in driving these innovations, ensuring that the benefits of AI and ML in fraud detection are realized across all sectors of the financial industry.

As financial fraud schemes continue to evolve, so too must the technologies designed to detect and prevent them. The continuous refinement of ML and DL models, coupled with advancements in data processing and analysis, sets the stage for a future where financial systems are not only more secure but also more resilient against the threats posed by fraud.

E. Conclusion

In reviewing the significant strides made in financial fraud detection through ML and DL, it becomes evident that these technologies have substantially advanced the capacity to identify and prevent fraudulent activities across diverse sectors.

The transition from traditional methodologies to more sophisticated models, including neural networks and hybrid systems, underscores a shift towards more dynamic, scalable, and efficient fraud detection mechanisms. This evolution is particularly vital in handling the complex and often imbalanced datasets typical in fraud scenarios, enhancing detection accuracy and operational efficiency.

Recent innovations, such as the integration of AI with blockchain and IoT technologies, promise further enhancements by offering more secure and transparent environments for fraud detection. However, challenges persist, particularly in the need for extensive labeled datasets, the opaque nature of some DL models, and the ongoing threat of model bias and ethical concerns in AI deployment.

Future directions should focus on refining these models to improve transparency and fairness while extending the reach of these technologies globally, especially in under-equipped regions. Continued collaborative efforts between academia and industry are essential for fostering innovations that address these challenges, ensuring the robustness and reliability of fraud detection systems.

Overall, the integration of ML and DL in financial fraud detection not only enhances the ability to combat fraud but also aligns with broader efforts to secure financial systems against increasingly sophisticated fraud schemes, thereby bolstering global economic stability and trust in financial institutions.

F. References

[1] Y. Alghofaili, A. Albattah, and M. A. Rassam, "A Financial Fraud Detection Model Based on LSTM Deep Learning Technique," Journal of Applied Security Research, vol. 15, no. 4, pp. 498–516, Oct. 2020, doi: 10.1080/19361610.2020.1815491.

[2] S. Rukhsar, M. J. Awan, U. Naseem, et al. (2023). Artificial intelligence based sentence level sentiment analysis of COVID-19. Computer Systems Science and Engineering, 47(1), 791-807.

[3] A. R. Khalid, N. Owoh, O. Uthmani, M. Ashawa, J. Osamor, and J. Adejoh, "Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach," Big Data and Cognitive Computing, vol. 8, no. 1, Jan. 2024, doi: 10.3390/bdcc8010006.

[4] S. Islam, M. M. Haque, and A. N. M. R. Karim, "A rule-based machine learning model for financial fraud detection," International Journal of Electrical and Computer Engineering, vol. 14, no. 1, pp. 759–771, Feb. 2024, doi: 10.11591/ijece.v14i1.pp759-771.

[5] K. Kapadiya et al., "Blockchain and AI-Empowered Healthcare Insurance Fraud Detection: An Analysis, Architecture, and Future Prospects," IEEE Access, vol. 10, pp. 79606–79627, 2022, doi: 10.1109/ACCESS.2022.3194569.

[6] M. A. Mohammed, A. Lakhan, K. H. Abdulkareem, et al. (2023). Homomorphic federated learning schemes enabled pedestrian and vehicle detection system. Internet of Things, 23, 100903.

[7] H. R. Abdulqadir, A. M. Abdulazeez, & D. A. Zebari. (2021). Data mining classification techniques for diabetes prediction. Qubahan Academic Journal, 1(2), 125-133.

[8] F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan, and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," IEEE Access, vol. 10, pp. 39700–39715, 2022, doi: 10.1109/ACCESS.2022.3166891.

[9] J. Homepage, O. J. Unogwu, and Y. Filali, "Wasit Journal of Computer and Mathematics Science Fraud Detection and Identification in Credit Card Based on Machine Learning Techniques", doi: 10.31185/wjcm.185.

[10] M. Maashi, B. Alabduallah, and F. Kouki, "Sustainable Financial Fraud Detection Using Garra Rufa Fish Optimization Algorithm with Ensemble Deep Learning," Sustainability (Switzerland), vol. 15, no. 18, Sep. 2023, doi: 10.3390/su151813301.

[11] S. Jiang, R. Dong, J. Wang, and M. Xia, "Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network," Systems, vol. 11, no. 6, Jun. 2023, doi: 10.3390/systems11060305.

[12] B. Lebichot, "Deep-learning domain adaptation techniques for credit cards fraud detection."

[13] T. Karthikeyan, M. Govindarajan, and V. Vijayakumar, "An effective fraud detection using competitive swarm optimization based deep neural network," Measurement: Sensors, vol. 27, Jun. 2023, doi: 10.1016/j.measen.2023.100793.

[14] H. Zhou, G. Sun, S. Fu, L. Wang, J. Hu, and Y. Gao, "Internet Financial Fraud Detection Based on a Distributed Big Data Approach with Node2vec," IEEE Access, vol. 9, pp. 43378–43386, 2021, doi: 10.1109/ACCESS.2021.3062467.

[15] M. N. Ashtiani and B. Raahemi, "Intelligent Fraud Detection in Financial Statements Using Machine Learning and Data Mining: A Systematic Literature Review," IEEE Access, vol. 10. Institute of Electrical and Electronics Engineers Inc., pp. 72504–72525, 2022. doi: 10.1109/ACCESS.2021.3096799.

[16] X. Mao, H. Sun, X. Zhu, and J. Li, "Financial fraud detection using the related-party transaction knowledge graph," in Procedia Computer Science, Elsevier B.V., 2021, pp. 733–740. doi: 10.1016/j.procs.2022.01.091.

[17] P. K. Kamuangu, "Journal of Economics, Finance and Accounting Studies A Review on Financial Fraud Detection using AI and Machine Learning," 2024, doi: 10.32996/jefas.

[18] A. T. El-Toukhy, M. M. Badr, M. M. E. A. Mahmoud, G. Srivastava, M. M. Fouda, and M. Alsabaan, "Electricity Theft Detection Using Deep Reinforcement Learning in Smart Power Grids," IEEE Access, vol. 11, pp. 59558–59574, 2023, doi: 10.1109/ACCESS.2023.3284681.

[19] J. K. Afriyie et al., "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," Decision Analytics Journal, vol. 6, Mar. 2023, doi: 10.1016/j.dajour.2023.100163.

[20] T. Ashfaq et al., "A Machine Learning and Blockchain Based Efficient Fraud Detection Mechanism," Sensors, vol. 22, no. 19, Oct. 2022, doi: 10.3390/s22197162.

[21] S. Misra, S. Thakur, M. Ghosh, and S. K. Saha, "An Autoencoder Based Model for Detecting Fraudulent Credit Card Transaction," in Procedia Computer Science, Elsevier B.V., 2020, pp. 254–262. doi: 10.1016/j.procs.2020.03.219.

[22] N. Nayyer, N. Javaid, M. Akbar, A. Aldegheishem, N. Alrajeh, and M. Jamil, "A New Framework for Fraud Detection in Bitcoin Transactions Through Ensemble Stacking Model in Smart Cities," IEEE Access, vol. 11, pp. 90916–90938, 2023, doi: 10.1109/ACCESS.2023.3308298.

[23] T. G.S., S. Dheeshjith, S. S. Iyengar, N. R. Sunitha, and P. Badrinath, "A hybrid and effective learning approach for Click Fraud detection," Machine Learning with Applications, vol. 3, p. 100016, Mar. 2021, doi: 10.1016/j.mlwa.2020.100016.

[24] S. Ryu, B. Jeon, H. Seo, M. Lee, J. W. Shin, and Y. Yu, "Development of deep autoencoder-based anomaly detection system for HANARO," Nuclear Engineering and Technology, vol. 55, no. 2, pp. 475–483, Feb. 2023, doi: 10.1016/j.net.2022.10.009.

[25] S. M. Abas, A. M. Abdulazeez, and D. Q. Zeebaree, "A YOLO and convolutional neural network for the detection and classification of leukocytes in leukemia," Indonesian Journal of Electrical Engineering and Computer Science, vol. 25, no. 1, pp. 200–213, Jan. 2022, doi: 10.11591/ijeecs.v25.i1.pp200-213.

[26] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," J Big Data, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40537-022-00573-8.

[27] C. Qiu, X. Wang, H. Yao, J. Du, F. R. Yu, and S. Guo, "Networking Integrated Cloud-Edge-End in IoT: A Blockchain-Assisted Collective Q-Learning Approach," IEEE Internet Things J, vol. 8, no. 16, pp. 12694–12704, Aug. 2021, doi: 10.1109/JIOT.2020.3007650.

[28] S. A. Korkmaz and F. Karataş, "Big Data: Controlling Fraud by Using Machine Learning Libraries on Spark," International Journal of Applied Mathematics, Electronics and Computers, vol. 6, no. 1, pp. 1–5, Mar. 2018, doi: 10.18100/ijamec.2018138629.

[29] M. Ravinder, V. Kulkarni, and R. Scholar, "Optimizing Anomaly Detection in Smart Grids with Modiied FDA and Dilated GRU-based Adaptive Residual RNN Optimizing Anomaly Detection in Smart Grids with Modified FDA and Dilated GRU-based Adaptive Residual RNN", doi: 10.21203/rs.3.rs-3869400/v1.

[30] D. Nandu Gawade, S. Dilip Turukmane, M. Rameshwar Bhagwat, and P. Madhav Kale, "Real-time Emotional Self-regulation and Stress Monitoring System." [Online]. Available: https://www.researchgate.net/publication/370058898

[31] G. C. Sekhar and R. Aruna, "A Novel Blockchain-Assisted Deep Learning Model For Secure Edge Intelligence in IoT Networks," Journal of The Institution of Engineers (India): Series C, Apr. 2024, doi: 10.1007/s40032-024-01048-w.

[32] L. N. CheSuh, R. Á. Fernández-Diaz, J. M. Alija-Perez, C. Benavides-Cuellar, and H. Alaiz-Moreton, "Improve quality of service for the Internet of Things using Blockchain & machine learning algorithms," Internet of Things (Netherlands), vol. 26, Jul. 2024, doi: 10.1016/j.iot.2024.101123.

[33] N. Omar, A. M. Abdulazeez, A. Sengur, and S. G. S. Al-Ali, "Fused faster RCNNs for efficient detection of the license plates," Indonesian Journal of Electrical

Engineering and Computer Science, vol. 19, no. 2, pp. 974–982, Aug. 2020, doi: 10.11591/ijeecs.v19.i2.pp974-982.

[34] M. Kaveh and M. S. Mesgari, "Application of Meta-Heuristic Algorithms for Training Neural Networks and Deep Learning Architectures: A Comprehensive Review," Neural Processing Letters, vol. 55, no. 4. Springer, pp. 4519–4622, Aug. 01, 2023. doi: 10.1007/s11063-022-11055-6.

[35] Y. Ding, W. Kang, J. Feng, B. Peng, and A. Yang, "Credit Card Fraud Detection Based on Improved Variational Autoencoder Generative Adversarial Network," IEEE Access, vol. 11, pp. 83680–83691, 2023, doi: 10.1109/ACCESS.2023.3302339.

[36] K. Ishida, Y. Takena, Y. Nota, R. Mochizuki, I. Matsumura, and G. Ohashi, "SA-PatchCore: Anomaly Detection in Dataset With Co-Occurrence Relationships Using Self-Attention," IEEE Access, vol. 11, pp. 3232–3240, 2023, doi: 10.1109/ACCESS.2023.3234745.

[37] T. R. Noviandy, G. M. Idroes, A. Maulana, I. Hardi, E. S. Ringga, and R. Idroes, "Credit Card Fraud Detection for Contemporary Financial Management Using XGBoost-Driven Machine Learning and Data Augmentation Techniques," Indatu Journal of Management and Accounting, vol. 1, no. 1, pp. 29–35, Sep. 2023, doi: 10.60084/ijma.v1i1.78.

[38] J. Debener, V. Heinke, and J. Kriebel, "Detecting insurance fraud using supervised and unsupervised machine learning," Journal of Risk and Insurance, vol. 90, no. 3, pp. 743–768, Sep. 2023, doi: 10.1111/jori.12427.

[39] Y. Yoo, J. Shin, and S. Kyeong, "Medicare Fraud Detection Using Graph Analysis: A Comparative Study of Machine Learning and Graph Neural Networks," IEEE Access, vol. 11, pp. 88278–88294, 2023, doi: 10.1109/ACCESS.2023.3305962.

[40] M. Rasheduzzaman Labu and M. Fahim Ahammed, "Next-Generation Cyber Threat Detection and Mitigation Strategies: A Focus on Artificial Intelligence and Machine Learning," 2024, doi: 10.32996/jcsts.

[41] M. Aljabri and R. M. A. Mohammad, "Click fraud detection for online advertising using machine learning," Egyptian Informatics Journal, vol. 24, no. 2, pp. 341–350, Jul. 2023, doi: 10.1016/j.eij.2023.05.006.

[42] H. Najadat, O. Altiti, A. A. Aqouleh, and M. Younes, "Credit Card Fraud Detection Based on Machine and Deep Learning," in 2020 11th International Conference on Information and Communication Systems, ICICS 2020, Institute of Electrical and Electronics Engineers Inc., Apr. 2020, pp. 204–208. doi: 10.1109/ICICS49469.2020.239524.

[43] M. Schneider and R. Brühl, "Disentangling the black box around CEO and financial information-based accounting fraud detection: machine learning-based evidence from publicly listed U.S. firms," Journal of Business Economics, vol. 93, no. 9, pp. 1591–1628, Nov. 2023, doi: 10.1007/s11573-023-01136-w.

[44] S. Agarwal, "An Intelligent Machine Learning Approach for Fraud Detection in Medical Claim Insurance: A Comprehensive Study," Scholars Journal of Engineering and Technology, vol. 11, no. 09, pp. 191–200, Sep. 2023, doi: 10.36347/sjet.2023.v11i09.003.

[45] A. Sina, "Open AI and its Impact on Fraud Detection in Financial Industry," Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), vol. 2, no. 3, pp. 263–281, Sep. 2024, doi: 10.60087/jklst.vol2.n3.p281.

[46] A. Salam et al., "Securing Smart Manufacturing by Integrating Anomaly Detection with Zero-Knowledge Proofs," IEEE Access, vol. 12, pp. 36346–36360, 2024, doi: 10.1109/ACCESS.2024.3373697.

[47] S. Dasari and R. Kaluri, "An Effective Classification of DDoS Attacks in a Distributed Network by Adopting Hierarchical Machine Learning and Hyperparameters Optimization Techniques," IEEE Access, vol. 12, pp. 10834–10845, 2024, doi: 10.1109/ACCESS.2024.3352281.

[48] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," Computer Science Review, vol. 40. Elsevier Ireland Ltd, May 01, 2021. doi: 10.1016/j.cosrev.2021.100402.

[49] F. R. Alzaabi and A. Mehmood, "A Review of Recent Advances, Challenges, and Opportunities in Malicious Insider Threat Detection Using Machine Learning Methods," IEEE Access, vol. 12, pp. 30907–30927, 2024, doi: 10.1109/ACCESS.2024.3369906.

[50] F. S. Alsubaei, A. A. Almazroi, and N. Ayub, "Enhancing Phishing Detection: A Novel Hybrid Deep Learning Framework for Cybercrime Forensics," IEEE Access, vol. 12, pp. 8373–8389, 2024, doi: 10.1109/ACCESS.2024.3351946.

[51] B. Dangsawang and S. Nuchitprasitchai, "A machine learning approach for detecting customs fraud through unstructured data analysis in social media," Decision Analytics Journal, vol. 10, Mar. 2024, doi: 10.1016/j.dajour.2024.100408.

[52] A. Wahid, M. Msahli, A. Bifet, and G. Memmi, "NFA: A neural factorization autoencoder based online telephony fraud detection," Digital Communications and Networks, vol. 10, no. 1, pp. 158–167, Feb. 2024, doi: 10.1016/j.dcan.2023.03.002.

[53] Z. Saad Rubaidi, B. Ben Ammar, and M. Ben Aouicha, "Vehicle Insurance Fraud Detection Based on Hybrid Approach for Data Augmentation," 2023. [Online]. Available: www.mirlabs.net/jias/index.html

[54] D. Jamil et al., "SUSTAINABLE FRAUD DETECTION IN GREEN FINANCE EMPOWERED WITH MACHINE LEARNING APPROACH," vol. 9, no. 1, pp. 1897–1914, 2024, doi: 10.33282/rr.vx9il.82.

[55] F. A. Nama and A. J. Obaid, "Financial Fraud Identification Using Deep Learning Techniques," Al-Salam Journal for Engineering and Technology, vol. 3, no. 1, pp. 141–147, Jan. 2024, doi: 10.55145/ajest.2024.03.01.012.

[56] A. Raza, K. Munir, M. S. Almutairi, and R. Sehar, "Novel Class Probability Features for Optimizing Network Attack Detection With Machine Learning," IEEE Access, vol. 11, pp. 98685–98694, 2023, doi: 10.1109/ACCESS.2023.3313596.

[57] M. A. Islam, M. A. Uddin, S. Aryal, and G. Stea, "An ensemble learning approach for anomaly detection in credit card data with imbalanced and overlapped classes," Journal of Information Security and Applications, vol. 78, Nov. 2023, doi: 10.1016/j.jisa.2023.103618.

[58] R. Jayaraj, A. Pushpalatha, K. Sangeetha, T. Kamaleshwar, S. Udhaya Shree, and D. Damodaran, "Intrusion detection based on phishing detection with machine learning," Measurement: Sensors, vol. 31, Feb. 2024, doi: 10.1016/j.measen.2023.101003.

[59] M. S. Cheong, M. C. Wu, and S. H. Huang, "Interpretable Stock Anomaly Detection Based on Spatio-Temporal Relation Networks with Genetic Algorithm," IEEE Access, vol. 9, pp. 68302–68319, 2021, doi: 10.1109/ACCESS.2021.3077067.

[60] C. G. Tekkali and K. Natarajan, "RDQN: ensemble of deep neural network with reinforcement learning in classification based on rough set theory for digital transactional fraud detection," Complex and Intelligent Systems, vol. 9, no. 5, pp. 5313–5332, Oct. 2023, doi: 10.1007/s40747-023-01016-4.

[61] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," J Big Data, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00541-8.

[62] R. Bin Sulaiman, V. Schetinin, and P. Sant, "Review of Machine Learning Approach on Credit Card Fraud Detection," Human-Centric Intelligent Systems, vol. 2, no. 1–2, pp. 55–68, Jun. 2022, doi: 10.1007/s44230-022-00004-0.

[63] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," IEEE Access, vol. 11, pp. 3034–3043, 2023, doi: 10.1109/ACCESS.2022.3232287.

[64] S. N. Kalid, K. C. Khor, K. H. Ng, and G. K. Tong, "Detecting Frauds and Payment Defaults on Credit Card Data Inherited with Imbalanced Class Distribution and Overlapping Class Problems: A Systematic Review," IEEE Access, vol. 12, pp. 23636–23652, 2024, doi: 10.1109/ACCESS.2024.3362831.

[65] A. A. Taha and S. J. Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine," IEEE Access, vol. 8, pp. 25579–25587, 2020, doi: 10.1109/ACCESS.2020.2971354.

[66] Z. Huang, H. Zheng, C. Li, and C. Che, "Application of Machine Learning-Based K-means Clustering for Financial Fraud Detection," 2024.

[67] F. Aslam, "Advancing Credit Card Fraud Detection: A Review of Machine Learning Algorithms and the Power of Light Gradient Boosting," American Journal of Computer Science and Technology, Feb. 2024, doi: 10.11648/ajcst.20240701.12.

[68] B. Omair and A. Alturki, "A Systematic Literature Review of Fraud Detection Metrics in Business Processes," IEEE Access, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 26893–26903, 2020. doi: 10.1109/ACCESS.2020.2971604.

[69] P. Craja, A. Kim, and S. Lessmann, "Deep learning for detecting financial statement fraud," Decis Support Syst, vol. 139, Dec. 2020, doi: 10.1016/j.dss.2020.113421.