Reviewing Machine Learning Techniques in Credit Card Fraud Detection

Ibtissam Medarhri[1][a], Mohamed Hosni[2][b], Mohamed Ettalhaoui[2], Zakaria Belhaj[1] and Rabie Zine[3][c]

[1]*MMCS Research Team, LMAID, ENSMR-Rabat, Morocco*

[2]*MOSI Research Team, LM2S3, ENSAM, Moulay Ismail University of Meknes, Meknes, Morocco*

[3]*School of Science and Engineering, Al Akhawayn University in Ifrane, Ifrane, Morocco*

Keywords: Credit Card Fraud, Machine Learning, Classification, Systematic Mapping Study.

Abstract: The growing use of credit cards for transactions has increased the risk of fraud, as fraudsters frequently attempt to exploit these transactions. Consequently, credit card companies need decision support systems that can automatically detect and manage fraudulent activities without human intervention, given the vast volume of daily transactions. Machine learning techniques have emerged as a powerful solution to address these challenges. This paper provides a comprehensive overview of the knowledge domain related to the application of machine learning techniques in combating credit card fraud. To achieve this, a review of published work in academic journals from 2018 to 2023 was conducted, encompassing 131 papers. The review classifies the studies based on eight key aspects: publication trends and venues, machine learning approaches and techniques, datasets, evaluation frameworks, balancing techniques, hyperparameter optimization, and tools used. The main findings reveal that the selected studies were published across various journal venues, employing both single and ensemble machine learning approaches. Decision trees were identified as the most frequently used technique. The studies utilized multiple datasets to build models for detecting credit card fraud and explored various preprocessing steps, including feature engineering (such as feature extraction, construction, and selection) and data balancing techniques. Python and its associated libraries were the most commonly used tools for implementing these models.

1 INTRODUCTION

The advancement of technology has significantly influenced the transition from traditional payment methods to online transactions (Mienye et al., 2023), (Taha and Malebary, 2020). Modern banking systems are now offering a wide array of payment options to enhance customer experience, including card payments, internet banking, and various e-services.

Globally, credit cards remain the most widely used payment method. According to the Nil Report (Report, 2023), there are 1,103 credit card issuers worldwide. In 2021, the combined purchase volume of the top 150 portfolios reached 12.695 trillion, reflecting a 9.4% increase compared to 2020.

While credit cards offer convenience for online purchases of goods and services, they also expose users to the risk of fraudulent transactions (Kim et al., 2019). In 2021 alone, 32.34 billion payment cards were compromised globally due to fraud (Report, 2023). Projections estimate that fraud-related losses will reach 408 billion over the next decade.

Current fraud detection systems predominantly rely on manually designed rules, which are often inefficient, subjective, and vulnerable to manipulation by fraudsters (Kim et al., 2019; Carcillo et al., 2018). As a result, there is a pressing need for automated detection systems. The growing adoption of electronic payment systems provides credit card issuers with extensive customer data, which can be leveraged to develop data-driven models that effectively detect fraud and minimize losses (Carcillo et al., 2018; Cheon et al., 2021; Pozzolo et al., 2018).

Machine Learning (ML) techniques have emerged

[a] https://orcid.org/0009-0003-0052-8702
[b] https://orcid.org/0000-0001-7336-4276
[c] https://orcid.org/0000-0002-0882-1327

Medarhri, I., Hosni, M., Ettalhaoui, M., Belhaj, Z. and Zine, R.
Reviewing Machine Learning Techniques in Credit Card Fraud Detection.
DOI: 10.5220/0013072500003838
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2024) - Volume 1: KDIR*, pages 179-187
ISBN: 978-989-758-716-0; ISSN: 2184-3228
Proceedings Copyright © 2024 by SCITEPRESS – Science and Technology Publications, Lda.

as a powerful tool for tackling credit card fraud (Pozzolo et al., 2018; Leevy et al., 2023; Salekshahrezaee et al., 2023). ML models, once deployed, can efficiently process large volumes of transactions in real-time, assuming the appropriate infrastructure is in place. The success of ML techniques has been demonstrated across various domains.

This paper presents a systematic mapping study aimed at gaining insights into the use of ML techniques in developing decision support systems for detecting fraudulent credit card transactions. The study examines key aspects, including publication trends and venues, ML approaches and techniques, datasets used for constructing Credit Card Fraud (CCF) models, evaluation frameworks, preprocessing techniques, hyperparameter optimization methods, and tools employed in model development.

The structure of the paper is as follows: Section 2 outlines the research protocol used in the study. Section 3 presents and discusses the findings for each mapping question. Finally, Section 4 concludes the paper and offers suggestions for future research.

2 RESEARCH PROTOCOL

This study aims to consolidate existing research on the application of ML in developing automated systems for credit card fraud management. To accomplish this, a systematic mapping study was conducted following the methodology outlined by (Petersen et al., 2008), which has been widely adopted in various research fields, including software engineering (Hosni and Idri, 2018), medical informatics (Hosni et al., 2019), and urban flood hazard mapping (El baida et al., 2024). The mapping process consists of several steps, which are described in detail in the following subsections.

2.1 Mapping Questions

The goal of this review is to provide a comprehensive understanding of how ML techniques, particularly classification methods, are utilized in the development of CCF systems. To fulfill this objective, we formulated eight research questions (MQs), each designed to explore specific aspects of ML application in CCF. Table 1 lists these MQs along with the motivations behind each question.

2.2 Search Strategy

This step aims to identify candidate papers relevant to the topic of this study. The primary sources of papers are digital libraries that index research published by leading publishers worldwide. For this study, we selected the Scopus digital library as our primary source of candidate papers. The initial task was to construct a search string to be used as input for the Scopus search engine.

The search string was formulated based on the authors' expertise and knowledge. The search query used was:

TITLE-ABS-KEY((fraud OR ”Fraud detection” OR ”Fraud Analytics”) AND (”credit card” OR ”card payment*” OR ”Transaction Fraud”) AND (”Machine learning”))

The searches were conducted on metadata of titles, abstracts, and keywords of research works between the years 2018 and 2023. We have limited our search to articles in peer-reviewed journals. We set this limitation to ensure that the papers selected have undergone a satisfactory peer-reviewing process and hence command a high level of academic integrity and reliability.

2.3 Study Selection

The pool of candidate papers obtained through the Scopus search needed further filtering based on predefined inclusion and exclusion criteria. This step was crucial to ensure that only relevant papers addressing our MQs were included. To maintain accuracy, three researchers independently performed the filtering process. A paper was included if it met at least one inclusion criterion and none of the exclusion criteria. If the decision was unclear based on the metadata, the researcher proceeded to read the full paper. The inclusion and exclusion criteria were as follows:

Inclusion Criteria:

- Papers that specifically focus on building credit card fraud detection systems using ML techniques.
- Papers that aim to enhance existing ML techniques for credit card fraud detection.
- Papers that compare different ML techniques in the context of credit card fraud detection.

Exclusion Criteria:

- Papers not written in English.
- Papers that do not utilize ML techniques for credit card fraud detection.
- Papers that focus on detecting fraudulent transactions unrelated to credit cards.

Table 1: Mapping Questions and their Motivations.

Mapping Questions	Motivations
Which journal venues are the primary targets for the use of ML techniques in credit card fraud detection? And what is the frequency of publication has changed over time?	To identify the specific journal venues where research related to ML techniques in credit card fraud detection is being published and discover the publication trend over time.
What are the ML approaches used in credit card fraud detection? Additionally, which specific ML techniques are commonly utilized?	To identify the various types of ML techniques used in CCFD systems and provide an enumeration of specific ML techniques that have been adopted in building these systems.
What are the main datasets used in credit card fraud detection?	To identify the prevalent datasets that researchers rely on when developing and evaluating CCFD systems.
What are the performance frameworks used to build and assess the credit card fraud detection model?	To identify the evaluation methods used to build the CCFD systems and enumerate the performance indicators used to assess the built models.
What techniques are used to handle the balancing problem in credit card fraud detection?	To identify the techniques used to handle the balancing problem present in CCF datasets.
What feature engineering stages have been investigated in the context of credit card fraud detection? Additionally, what are the techniques that have been used in each of these stages?	To identify the feature engineering stages that have been treated in literature. Furthermore, enumerate the techniques used in each of the identified stages.
What are the optimization techniques used to fine-tune the hyperparameters of the ML techniques in credit card fraud detection systems?	To identify the optimization techniques used to fine-tune the hyperparameters of the ML techniques in credit card fraud detection.
What tools are used to build credit card fraud detection models?	To identify the tools used to build credit card fraud detection models.

2.4 Data Extraction and Synthesis

After selecting the papers relevant to our MQs, data extraction was performed independently by three researchers. The extracted data were systematically recorded in detailed forms, ensuring alignment with each MQ.

Following a comprehensive review of the extracted data, synthesis was conducted by summarizing and aggregating the findings for each MQ from all selected papers. Two synthesis methods were employed: narrative synthesis and the counting method, which allowed for the consistent tabulation of data in line with the MQs. Visualization tools, such as bar charts and pie charts, were used to present the aggregated data.

3 RESULTS AND DISCUSSION

This section presents and discusses the results obtained from the mapping study, organized according to the research questions listed in Table 1.

3.1 Results Overview

A total of 790 candidate papers were retrieved through the automatic search in the Scopus database using the search string specified in Section 2.2. The search was restricted in two ways: first, by time frame, including only papers published between 2018 and 2023, and second, by selecting only journal articles. The search was conducted on June 24, 2024. The primary reason for limiting the search to 2023 is to facilitate the replication of the search results, as the likelihood of additional papers being indexed for that year is minimal. In contrast, selecting an ongoing year could pose challenges since the indexing process for papers published within the same year may take time to complete.

Following the study selection process and the application of inclusion and exclusion criteria, 131 papers were selected. Relevant information was then extracted from these papers to address the research questions (MQs). It is worth noting that both the selection and data extraction processes were performed independently by three researchers. Additionally, not all 131 papers provided answers to all the research questions. Details of the selected papers and extracted data are available upon request.

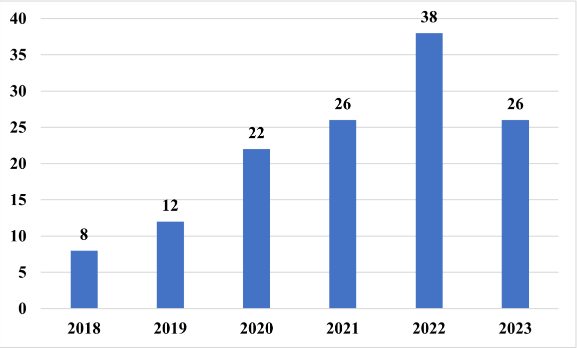

Figure 1: Publication Trends over Time.

3.2 Publication Venues and Trends (MQ1)

This review identified 86 different venues where the 131 selected papers were published. The IEEE Access journal had the highest number of publications, with 13 papers, followed by the Journal of Theoretical and Applied Information Technology with five publications and the Journal of Big Data with four. Seven journals published three papers each, while thirteen journals published two papers each. Additionally, 63 venues published only one paper each. Table 2 lists the main sources that published more than three papers.

Regarding publication trends, an upward trajectory in the number of publications was observed over time. It is important to note that only papers published in journals over the last five years were included in this review. The highest number of publications occurred in 2022, with 38 papers published across 28 different venues. IEEE Access led with four papers, followed by seven journals that published two papers each, while the remaining papers were distributed among 20 other journals, each publishing one paper. Figure 1 illustrates the publication trends over the search period.

3.3 Machine Learning: Approaches and Techniques (MQ2)

The objective of the MQ2 is to identify the most prevalent ML approaches used by researchers and to catalog the specific ML techniques employed in the selected studies.

Figure 2 illustrates the distribution of ML approaches used in the reviewed papers. The findings show that 39% of the selected studies (51 out of 131 papers) focused exclusively on single ML approaches. Meanwhile, 29% of the papers (39 out of 131) explored ensemble ML approaches alone. Notably, 32% of the papers (42 out of 131) investigated both single and ensemble approaches.

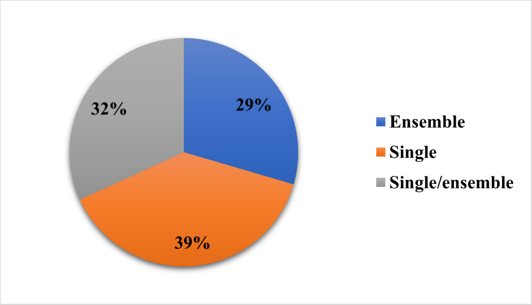

Figure 2: Publication Trends over Time.

Table 3 provides a comprehensive list of ML techniques that have been applied in developing decision support systems for detecting fraudulent credit card transactions (CCFD). The review identified 11 single classification techniques commonly explored in CCFD literature. Among these, Decision Tree (DT) was the most frequently used technique, appearing in 82 instances. Artificial Neural Networks (ANN) were investigated 67 times, while Regression techniques were utilized 47 times. Support Vector Machines (SVM) were employed in 32 instances. Notably, four techniques were each used only once.

Out of the 131 selected papers, 60 focused on investigating a single ML technique, and nine papers examined two ML techniques. The study that explored the highest number of ML techniques, totaling 31, was (Randhawa et al., 2018).

Ensemble methods were explored in 113 instances within the selected studies. The primary type of ensemble investigated was homogeneous, particularly the combination of a single base technique with a meta ensemble technique. Among the meta ensemble techniques, Boosting was the most commonly used, with XGBoost being the most extensively studied, appearing in 22 cases. Other meta ensemble techniques, such as Random Subspace and Bagging, were also explored. Additionally, heterogeneous ensembles were investigated in the selected studies (Baker, 2022).

3.4 Datasets Used (MQ3)

The construction of CCF models primarily relies on historical transaction data. This MQ aims to identify and catalog the datasets used in the selected studies for building CCF models. A total of 29 different datasets were identified across the selected studies. Table 4 lists the datasets that were utilized more than four times. Notably, the ”Credit Card Fraud Dataset,” containing 284,807 records, was the most frequently

Table 2: Publication Venues.

Journal	Number
IEEE Access	13
Journal of Theoretical and Applied Information Technology	5
Journal of Big Data	4
Multimedia Tools and Applications	3
International Journal of Intelligent Engineering and Systems	3
International Journal of Interactive Mobile Technologies	3
International Journal on Recent and Innovation Trends in Computing and Communication	3
Applied Sciences (Switzerland)	3
Electronics (Switzerland)	3
Mathematics	3

Table 3: ML techniques used in the Selected Studies.

Technique	Number
Ensemble	113
DT	82
ANN	67
Regression	47
SVM	32
KNN	25
NB	23
Rule	3
Independent component analysis	1
K-means	1
Local Outlier Factor	1
PCA	1

used, appearing in 85 out of the 131 selected papers. This dataset is publicly available on the Kaggle platform. Additionally, 16 papers employed more than one dataset, with the maximum number of datasets used in a single study being three, as reported in three papers (Arora et al., 2020; de Zarzà et al., 2023; Zhu et al., 2020).

The review also identified several studies that utilized private datasets, including those collected from organizations in China (Zheng et al., 2020; Li et al., 2021b), various European countries (Marco et al., 2022), and financial institutions in South Korea (Kim et al., 2019), among others. It is important to note that most of the datasets used suffered from the problem of data imbalance, where the fraudulent class was significantly underrepresented compared to the non-fraudulent class.

3.5 Evaluation Framework: Evaluation Methods and Performance Metrics (MQ4)

The MQ4 aims to identify the evaluation frameworks used to assess CCF models in the selected studies. It specifically focuses on the evaluation methods employed to develop CCF models and the performance indicators used to measure their predictive capabilities. The review identified 38 different performance criteria. Table 5 lists the nine performance indicators that were used more than ten times to evaluate the predictive capabilities of the ML techniques applied in the selected studies.

The most frequently used performance criterion was Sensitivity, appearing in 115 instances. Precision and Accuracy were used 95 and 89 times, respectively. The F1-score and ROC AUC were also commonly adopted, appearing 79 and 69 times, respectively. One of the selected studies utilized ten performance indicators to assess the proposed models. Notably, 121 out of the 131 selected papers employed more than one performance criterion to evaluate their models.

Regarding the validation techniques used in building the ML models, Table 6 lists the different validation approaches investigated in the literature along with their frequency of use. A total of four validation approaches were identified. The Holdout validation technique was the most frequently used, appearing in 61 research papers. It was followed by the K-fold cross-validation technique, employed in 42 papers. Among these, 10-fold cross-validation was the most common, appearing in 21 papers, followed by 5-fold cross-validation. Notably, four papers did not specify the number of folds used. The stratified K-fold and cross-validation techniques were each adopted in six papers. It is also worth noting that some papers did not provide details about the validation technique used to develop their models.

3.6 Handling Balancing Problem (MQ5)

This MQ aims to explore how the issue of imbalanced datasets is addressed in the selected studies. Imbalanced datasets, where the number of fraudulent trans-

Table 4: Datasets used in the selected studies.

Dataset	Number
Credit Card Fraud Detection Dataset	85
Default of Credit Card Clients Dataset	7
Vesta IEEE-CIS	5
Financial company in China	5
BankSim	4
Generated Dataset	4
Dataset emerges from Kaggle	4
cc Fraud dataset	4
UCSD-FICO dataset	4

Table 5: Performance indicators used in the selected studies.

Performance Criterion	Number
Sensitivity	115
Precision	95
Accuracy	89
F1-score	79
AUC	69
Specificity	41
MCC	17
AUC-PR	15
False Positive Rate	15

Table 6: Validation techniques used in the selected studies.

Validation techniques	K	Number
Stratified	5 fold	3
	10 fold	3
K-cross validation	K-fold	4
	2 fold	1
	3 fold	1
	4 fold	1
	5 fold	13
	10 fold	21
	15 fold	1
Holdout		61
Cross validation		6

actions is significantly lower than that of legitimate transactions, pose challenges in training ML models effectively. Table 7 lists the balancing techniques that were used more than three times to handle class imbalance in the selected papers. A total of 32 techniques were identified.

The most widely adopted technique was SMOTE (Synthetic Minority Over-sampling Technique), which was used in 24% of the selected papers (31 out of 131). Following SMOTE, Random Under Sampling, Under Sampling, and Over Sampling techniques were utilized in 12, 11, and 10 papers, respectively. It is worth noting that four papers addressed the class imbalance problem without explicitly specifying the technique used (Bakhtiari et al., 2023), (Sadgali et al., 2021; Rakhshaninejad et al., 2022; Trisanto, 2021).

Table 7: Imbalanced techniques used in the selected studies.

Technique	Number
SMOTE	31
Random Under sampling	12
Under Sampling	11
Over Sampling	10
SMOTE-Edited Nearest Neighbors	7
Random Oversampling	5
SMOTE-Tomek	4
Addressed	4
Borderline SMOTE	3
Near Miss	3

Table 8: Feature Engineering aspects investigated in the selected studies.

Aspect	Number
Extraction	16
Feature Importance	4
Feature selection	41
Feature Construction	1

3.7 Feature Engineering: Steps Investigated, and Techniques Used (MQ6)

This MQ aims to explore the feature engineering approaches investigated by researchers in the selected studies and to identify the techniques employed at each step. Out of the 131 selected papers, 44 considered feature engineering as a crucial preprocessing step. Four key aspects of feature engineering were examined: feature construction, extraction, importance, and selection.

Among these aspects, feature selection was the

Table 9: Feature Extraction, Construction and Importance techniques used in the selected studies.

Extraction		Construction		Importance	
PCA	10	Feature Construction	1	XGBoost	2
Auto Encoder	4			LightGBM	1
Convolutional Neural Network	1			Shapley addictive explanations	1
Linear Discriminant Analysis	1				

Table 10: Feature Selection Techniques investigated in the selected studies.

Filter Techniques		Wrapper Techniques	
Correlation	10	Genetic Algorithm	2
Information Gain	5	Recursive Feature Elimination	2
Random Forest	3	Stepwise	2
Chi2	1	Rock Hyrax Swarm Optimization	1
Correlation based Feature Selection	1	SVM Recursive Elimination	1
Decision Tree	1	Quantum Algorithm Feature Selection by Q-SVM	1
Degree Centrality	1		
Distance based Feature Selection	1		
Entropy	1		
Extra Tree Ensemble	1		
Gain Ration	1		
LASSO	1		
Mutual Information	1		
ReliefF	1		
Factorial Analysis of Mixed Data	1		
Rough set	1		
standardized murals with ANOVA F-values	1		

Table 11: Hyperparameters Optimization techniques used in the selected studies.

Optimization technique	Number
Grid Search	27
Adam	9
Given	7
Bayesian	4
Genetic Algorithm	3
Particle Swarm Optimization	3
Randomized Search CV	2
Default Parameters	2
Differential Evolution Algorithm	2
Firefly Algorithm	2

most extensively studied, appearing in 41 experiments. Feature extraction was explored in 16 experiments, as detailed in Table 8.

Four feature extraction techniques were identified, as listed in Table 9. The most commonly used technique was Principal Component Analysis (PCA), which appeared in 10 instances. This was followed by the Auto Encoder technique, used four times. Regarding feature construction, only one study specifically focused on this aspect, utilizing both domain knowledge and statistical methods to create new features (Wu et al., 2019). For feature importance, three techniques were employed: XGBoost was used twice, while LightGBM and the Shapley Additive Explanations (SHAP) model were each used once.

Table 12: ML tools used in the selected papers.

Tool	Number
Python	71
Weka	11
MATLAB	4
Java	4
R	3
LibSVM	1
Orange	1
RapidMiner	1
SAS E-miner	1

Regarding feature selection techniques, as detailed in Table 10, this review identified two main categories: filter and wrapper techniques. Among the filter techniques, 17 different methods were used across the experiments in the selected papers. The most frequently employed filter technique was the correlation coefficient, such as Pearson correlation, which was used in 10 experiments. Information Gain and Random Forest were utilized in 5 and 3 experiments, respectively, while the remaining 14 techniques were each explored once.

For wrapper techniques, six methods were identified in the selected studies. The Genetic Algorithm, Recursive Feature Elimination, and Stepwise techniques were each explored twice, while the other three techniques were used once.

3.8 Hyperparameters Optimization Techniques (MQ7)

Hyperparameter optimization is crucial for enhancing the performance and generalization ability of ML models. This question aims to identify the hyperparameter optimization techniques employed in the selected studies.

In this review, 20 different optimization techniques were identified, used to fine-tune the hyperparameters of ML models. These techniques are listed in Table 11. Notably, Grid Search was the most frequently adopted optimization method, appearing in 27 research papers. The Adam optimizer was explored in 9 papers. Additionally, seven papers explicitly listed the parameter values of their employed ML techniques, while two papers used the default parameters provided by the tools used.

It is important to highlight that only 57 out of the 131 selected papers considered the hyperparameter optimization step. Moreover, seven studies employed multiple optimization techniques (Zhu et al., 2020; Li et al., 2021b; Tayebi and El, 2022; Li et al., 2021a; Yara et al., 2020; Grossi et al., 2022; Sharma et al., 2021). The study with the most comprehensive exploration of optimization techniques investigated seven different methods (Tayebi and El, 2022).

3.9 ML Tools (MQ8)

This question aims to identify the tools used to build decision support systems for detecting fraudulent credit card transactions. Table 12 provides a list of the nine identified tools.

The Python programming language was the most widely used, appearing in 71 papers. The Weka tool was utilized in 11 papers, while MATLAB and Java were each employed in four papers. Additionally, four tools were used in only one paper each.

The identified tools can be categorized into two groups: those with a **user interface**, such as RapidMiner, Orange, SAS E-miner, and Weka, and those that provide a **programming environment**, such as MATLAB, Java, R, Python, and the Weka API.

4 CONCLUSIONS AND FUTURE WORK

This paper presents a systematic mapping study that structures the body of knowledge on the use of ML techniques in developing decision support systems for detecting fraudulent credit card transactions. The study reviewed papers published in journal venues indexed in the Scopus database between 2018 and 2023. After applying the study selection process, including specific inclusion and exclusion criteria, 131 papers were selected to address eight mapping questions. The main findings related to each mapping question, as outlined in Table 1, are summarized below:

- The selected papers were published across 86 different journal venues.
- Both single ML approaches and ensemble approaches were investigated, with single ML approaches being the most prevalent.
- A total of 29 different datasets were utilized to build credit card fraud detection systems.
- Various performance indicators were used to evaluate the predictive capabilities of the models, with the Holdout validation technique being the most frequently employed.
- A total of 32 balancing techniques were identified, with SMOTE being the most commonly used method.
- Feature extraction, construction, and selection steps were explored in the selected studies.
- Only 27 studies optimized the hyperparameter settings of the ML techniques used.
- Nine tools were identified for building credit card fraud detection systems in the selected studies.

Future research directions could include exploring the construction and effectiveness of ensemble techniques in credit card fraud detection systems. Another promising area of investigation is identifying the most effective ML models for distinguishing between fraudulent and legitimate transactions, which could be systematically explored through a comprehensive literature review.

REFERENCES

Arora, V., Leekha, R. S., Lee, K., and Kataria, A. (2020). Facilitating user authorization from imbalanced data logs of credit cards using artificial intelligence. *Mobile Information Systems*, 2020(1):8885269.

Baker, M. R. (2022). Ensemble learning with supervised machine learning models to predict credit card fraud transactions.

Bakhtiari, S., Nasiri, Z., and Vahidi, J. (2023). Credit card fraud detection using ensemble data mining methods. *Multimedia Tools and Applications*, 82(19):29057–29075.

Carcillo, F., Dal Pozzolo, A., Le Borgne, Y.-A., Caelen, O., Mazzer, Y., and Bontempi, G. (2018). Scarff: a scalable framework for streaming credit card fraud detection with spark. *Information fusion*, 41:182–194.

Cheon, M.-j., Lee, D., Joo, H. S., and Lee, O. (2021). Deep learning based hybrid approach of detecting fraudulent transactions. *Journal of Theoretical and Applied Information Technology*, 99(16):4044–4054.

de Zarzà, I., de Curtò, J., and Calafate, C. T. (2023). Optimizing neural networks for imbalanced data. *Electronics*, 12(12):2674.

El baida, M., Hosni, M., Boushaba, F., Chourak, M., et al. (2024). A systematic literature review on classification machine learning for urban flood hazard mapping. *Water Resources Management*, pages 1–42.

Grossi, M., Ibrahim, N., Radescu, V., Loredo, R., Voigt, K., and Altrock, C. V. O. N. (2022). Mixed quantum – classical method for fraud detection with quantum feature selection. *IEEE Trans. Quantum Eng.*, 3(October):1–12.

Hosni, M., Carrillo-de Gea, J. M., Idri, A., Fernández-Alemán, J. L., and García-Berná, J. A. (2019). Using ensemble classification methods in lung cancer disease. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1367–1370. IEEE.

Hosni, M. and Idri, A. (2018). Software development effort estimation using feature selection techniques. In *New trends in intelligent software methodologies, tools and techniques*, pages 439–452. IOS Press.

Kim, E. et al. (2019). Champion-challenger analysis for credit card fraud detection: Hybrid ensemble and deep learning. *Expert Syst. Appl.*, 128:214–224.

Leevy, J. L., Johnson, J. M., Hancock, J., and Khoshgoftaar, T. M. (2023). Threshold optimization and random undersampling for imbalanced credit card data. *J. Big Data*.

Li, C., Ding, N., Zhai, Y., and Dong, H. (2021a). Comparative study on credit card fraud detection based on different support vector machines. *Intelligent Data Analysis*, 25(1):105–119.

Li, Z., Huang, M., Liu, G., and Jiang, C. (2021b). A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection. *Expert Syst. Appl.*, 175(February):114750.

Marco, G. et al. (2022). The role of diversity and ensemble learning in credit card fraud detection. *Adv. Data Anal. Classif.*

Mienye, I. D., Sun, Y., and Member, S. (2023). A deep learning ensemble with data resampling for credit card fraud detection. *IEEE Access*, 11(February):30628–30638.

Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. (2008). Systematic mapping studies in software engineering. In *12Th International Conference on Evaluation and Assessment in Software Engineering*, page 10.

Pozzolo, A. D., Boracchi, G., Caelen, O., and Alippi, C. (2018). Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE Trans. Neural Networks Learn. Syst.*, 29(8):3784–3797.

Rakhshaninejad, M., Fathian, M., Amiri, B., and Yazdanjue, N. (2022). An ensemble-based credit card fraud detection algorithm using an efficient voting strategy. *The Computer Journal*, 65(8):1998–2015.

Randhawa, K., Loo, C. H. U. K., and Member, S. (2018). Credit card fraud detection using adaboost and majority voting. *IEEE Access*, 6:14277–14284.

Report, N. (October 2023). The world's top card issuers and merchant acquirers.

Sadgali, I., Sael, N., and Benabbou, F. (2021). Human behavior scoring in credit card fraud detection. *IAES International Journal of Artificial Intelligence*, 10(3):698.

Salekshahrezaee, Z., Leevy, J. L., and Khoshgoftaar, T. M. (2023). The effect of feature extraction and data sampling on credit card fraud detection. *J. Big Data*.

Sharma, P., Banerjee, S., Tiwari, D., and Patni, J. C. (2021). Machine learning model for credit card fraud detection- a comparative analysis. *The International Arab Journal of Information Technology*, 18(6):789–796.

Taha, A. A. and Malebary, S. J. (2020). An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. *IEEE access*, 8:25579–25587.

Tayebi, M. and El, S. (2022). Performance analysis of metaheuristics based hyperparameters optimization for fraud transactions detection. *Evol. Intell.*, page 0123456789.

Trisanto, D. (2021). Modified focal loss in imbalanced xgboost for credit card fraud detection. *Int. J. Ind. Eng. Syst.*, 14(4):350–358.

Wu, Y., Xu, Y., and Li, J. (2019). Feature construction for fraudulent credit card cash-out detection. *Decis. Support Syst.*, page 113155.

Yara, A., Albatul, A., and A, R. M. (2020). A financial fraud detection model based on lstm deep learning technique. *J. Appl. Secur. Res.*, 0(0):1–19.

Zheng, L., Liu, G., Yan, C., Jiang, C., Zhou, M., and Li, M. (2020). Improved tradaboost and its application to transaction fraud detection. *IEEE Transactions on Computational Social Systems*, 7(5):1304–1316.

Zhu, H., Liu, G., Zhou, M., Xie, Y., and Abusorrah, A. (2020). Optimizing weighted extreme learning machines for imbalanced classification and application to credit card fraud detection. *Neurocomputing*, 407:50–62.