IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 December 2023, accepted 28 January 2024, date of publication 6 February 2024, date of current version 16 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3362831

SURVEY

Detecting Frauds and Payment Defaults on Credit Card Data Inherited With Imbalanced Class Distribution and Overlapping Class Problems: A Systematic Review

SURAYA NURAIN KALID[1], KOK-CHIN KHOR[2], KENG-HOONG NG[2], AND GEE-KOK TONG[1]

[1]Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Selangor 63100, Malaysia
[2]Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Bandar Sungai Long, Kajang 43000, Malaysia

Corresponding author: Kok-Chin Khor (kckhor@utar.edu.my)

This work was supported by Multimedia University Malaysia.

ABSTRACT Credit card payments are one popular e-payment option apart from cash payments. Recent reports show that credit card fraud and payment defaults are increasing annually and are alarming. Thus, researchers have attempted various machine learning techniques to address these two challenges. However, they are challenged to mitigate the two major problems inherited in credit card data: (i) imbalanced class distribution and (ii) overlapping classes. Mitigating these problems shall effectively detect credit card frauds and payment defaults, thus benefiting card issuers and holders. Hence, this paper aims to develop a systematic review using PRISMA to identify and compare various credit card datasets, machine learning techniques, and evaluation metrics. Subsequently, we provide recommendations for handling these two problems. We extracted research papers from 2016 to 2023 from ScienceDirect, Springer, Association and Computing Machinery (ACM), and IEEE databases. The papers shall be included if written in English and published in peer-reviewed and indexed journals or conference proceedings. Finally, 87 papers were selected based on the eligibility criteria. Based on our findings, the European and Taiwan datasets are widely used in the research community. However, most researchers focus on tackling imbalanced class distribution rather than two problems together. We recommended to the research community the application of deep learning, ensemble learning, and sampling methods to effectively detect fraud and payment defaults on credit card datasets that inherit the two problems. In evaluating the machine learning algorithms, we recommend using metrics that can separately evaluate the algorithms' performance in detecting frauds/payment defaults and normal transactions.

INDEX TERMS PRISMA, credit card fraud, payment default, imbalanced class distribution, overlapping classes.

I. INTRODUCTION

Frauds and payment defaults on credit cards are the two main challenges issuers face. The mainstream media reported that the past COVID-19 pandemic has triggered a rise in card fraud cases when it accelerates the digitalization of economies [1], [2]. Fraudsters take advantage of various situations to execute their fraudulent activities, and card fraud activities have already caused a great loss globally. According to the Nilson Newsletter of December 2022, the losses were approximately \$32.34 billion worldwide in 2021 [3].

In Malaysia, using credit cards for e-payment transactions has increased by 28.7% from 2021 to 2022 [4]. According to a news report, Malaysian credit card holders have yet to pay RM35.89 billion in outstanding balance in July 2022 [5]. Credit card holders may get interest charges for late payments, collection calls and visits, black marks on the credit report, and even bankruptcy, should they fail to repay (payment default) [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Yudong Zhang.

It is reported that credit card companies such as Visa are looking into artificial intelligence (AI) technology to detect fraud cases. Financial institutions also train machine learning models using financial data to identify credit card holders prone to payment default based on their repayment delinquency.

Researchers have also made great efforts to deal with credit card fraud and payment defaults. However, they often face difficulties dealing with credit card data as it usually suffers from imbalanced and overlapping class problems. The imbalanced class problem occurs when the number of samples in a majority class (normal transactions) is much greater than in a minority class (frauds/payment defaults). The data involved is even more complicated when the minority class samples have similar characteristics to the majority class samples, thus causing the problem of overlapping classes. As a result, general machine learning algorithms learn the minority class ineffectively. Further, the algorithms bias the majority class to maximize the detection accuracy. Thus, the algorithms produce low detection rates for credit card fraud and payment defaults.

In this study, we aim to review credit card datasets, their inherited problems, machine learning techniques used to tackle the problems and evaluation metrics used to measure the techniques in credit card research. We shall recommend effective machine learning techniques to handle imbalanced class distribution and overlapping class problems in credit datasets, as well as suitable evaluation metrics for evaluating the techniques' performance.

II. RESEARCH QUESTIONS

The systematic review responds to the following research questions. We shall answer the questions in Section IV Results.

1. Which credit card datasets are widely used for fraud and payment default detection in the research community?
2. How frequently are the two problems, imbalanced class distribution and payment defaults, being addressed in credit card research?
3. Which machine learning techniques have been applied to address these two problems?
4. What metrics are appropriate to evaluate the performance of the machine learning techniques?
5. What is the highest detection result achieved so far?
6. Which machine learning techniques produce the promising detection result?

III. METHODS

This systematic review was conducted based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [7]. PRISMA is the preferred reporting item for systematic review in many research communities.

A. ELIGIBILITY CRITERIA

For papers to be included in this systematic review, the researchers must report machine learning techniques for tackling the imbalanced class distribution and overlapping class problems in credit card datasets. The papers shall report original research related to the credit card, irrespective of the maturity level of each published work. They must be written in English and published in peer-reviewed and indexed journals or conference proceedings.

B. INFORMATION SOURCES

We retrieved papers from the online bibliographic databases, i.e., ScienceDirect, Springer, ACM, and IEEE. Fig. 1 presents the PRISMA process for selecting research papers to answer the research questions.

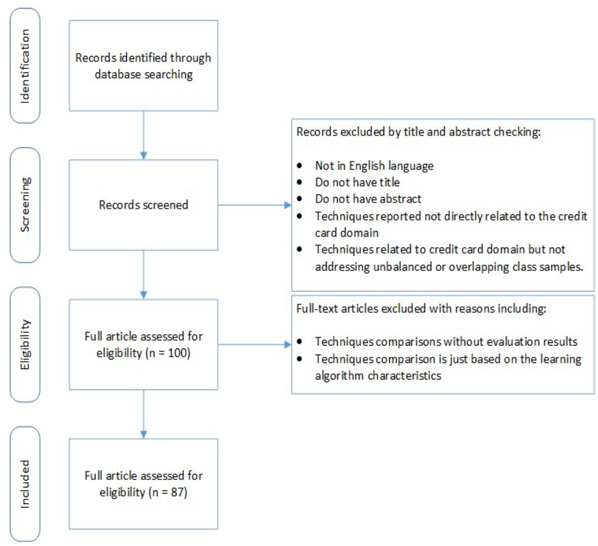

FIGURE 1. The PRISMA process for paper collection and analysis.

C. SEARCH STRATEGY

We searched the papers in the ScienceDirect, Springer, ACM, and IEEE databases from 2016 to 2023. The search strings were: (Credit Card OR Credit Card Fraud OR Credit Card Payment Default) AND (Imbalanced OR Unbalanced OR Rare Class OR Minority) OR (Overlapping). To define the search strings, we used the PICO criteria: problem (P), intervention (I), comparison (C), and outcome (O) [8]. Problem (P) shall represent the problems that we are looking into: (i) imbalanced class distribution and (ii) overlapping classes in credit card datasets. Interventions (I) are the machine learning techniques for solving problems. A comparison (C) was conducted by comparing different machine learning techniques used by researchers in addressing the problems. The outcomes (O) analyze techniques based on the appropriate evaluation metric.

D. STUDY SELECTION

We selected papers from peer-reviewed journals or conference proceedings related to the credit card domain in the screening. We checked the titles and abstracts of the papers and excluded papers without titles, abstracts and those that were not written in English. We also excluded papers that did not address the problems of imbalanced class distribution and overlapping classes. In the second screening, we read the full-text papers. The papers shall be excluded if they do not provide experimental results. After screening, we consolidated 87 papers for this systematic review and listed them in Appendix A.

E. DATA EXTRACTION

In this phase, we extracted six variables to answer the six research questions (refer to Section II). The variables include the widely used credit card datasets, the dataset problems, the machine learning techniques used to address the problems, evaluation metrics to measure the techniques' performance, the detection results and the effective machine learning techniques for detection.

IV. RESULTS

To answer the research questions, we summarized and presented the variables in the following figures and tables.

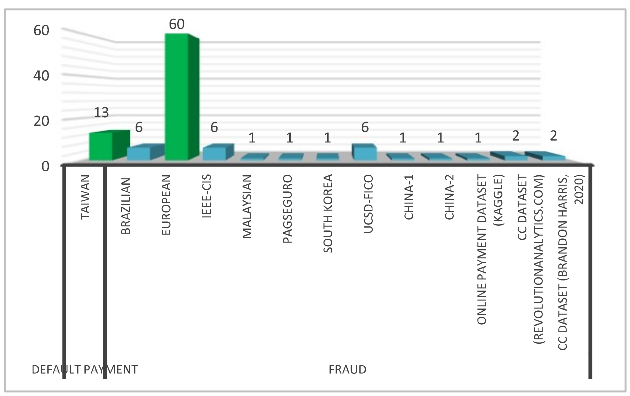

FIGURE 2. Frequency of the datasets used in the selected studies.

A. THE WIDELY USED CREDIT CARD DATASETS

We reviewed credit card fraud and payment default datasets employed by these 87 papers. Researchers may state the usage of more than one dataset in their papers. Fig. 2 demonstrates the usage frequency of each dataset. The descriptions of the datasets are shown in Table 1.

B. THE DATASETS' PROBLEMS

Dealing with credit card datasets is challenging for many researchers because of the two inherited dataset problems, i.e., imbalanced class distribution and overlapping class samples. Fig. 3 demonstrates the frequency of these problems being addressed in the selected papers. All the papers address the imbalanced class distribution problem based on the analysis. However, only 15 papers address the overlapping class problem together with the imbalanced class distribution problem. None address the overlapping class problem alone.

C. MACHINE LEARNING TECHNIQUES USED

Researchers used various machine learning techniques to address the two problems. Fig. 4 shows the frequency of these techniques being used for handling these two problems. Note that the researchers usually used multiple techniques in the same study to address the problems. Thus, the total frequency of techniques used is more than 87, the total number

TABLE 1. Summary of datasets used in the selected papers.

Dataset	Total Rec.	Description	% Normal	% F/PD**
Brazilian [22]	41,647	Real-time credit card data from a Brazilian bank.	96.26	3.74
CC Dataset (Brandon Harris, 2020) [68]	1,048,575	contains users' credit card transactions from January 1, 2019 to March 3, 2020.	99.43	0.57
CC Dataset (Revolution Analytics.com) [70]	10,000,000	Contains a list of credit card transactions	94.04	5.96
China-1 [25]*	2,258,036	Credit card fraud data from a financial company in China.	99.98	0.02
China-2 [45]*	153,685	Credit card fraud data from a financial institution in China.	98.50	1.50
European [96]	284,807	Credit card transactions of European cardholders in September 2013.	99.83	0.17
IEEE-CIS [97]	590,540	The fraud detection dataset was provided by Vesta for competition purposes.	99.97	0.03
Malaysian [15]*	287,224	Cardholders' data from the South-East Asia region from February to April 2017	99.96	0.04
Online Payment Dataset (Kaggle) [68]	1,048,575	an online payment dataset where transactions made on an online payment system	99.89	0.11
PagSeguro [16]	903,801	Online transaction data by PagSeguro in Brazil.	98.20	1.80
South Korea [39]*	11,025,756	14 months of transactions collected by a card issuer company in South Korea.	99.95	0.05
Taiwan [98]	30,000	Bill statements of credit card clients in Taiwan from April 2005 to September 2005.	78.00	22.00
UCSD-FICO [99]	94,683	UCSD-FICO data mining contest 2009 credit card dataset, an actual dataset of e-commerce transactions.	98.00	2.00

**The dataset source is not specified.*
***F/PD – Fraud/Payment Default*

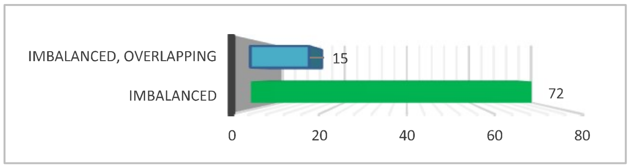

FIGURE 3. Frequency of the problems being addressed in the selected papers.

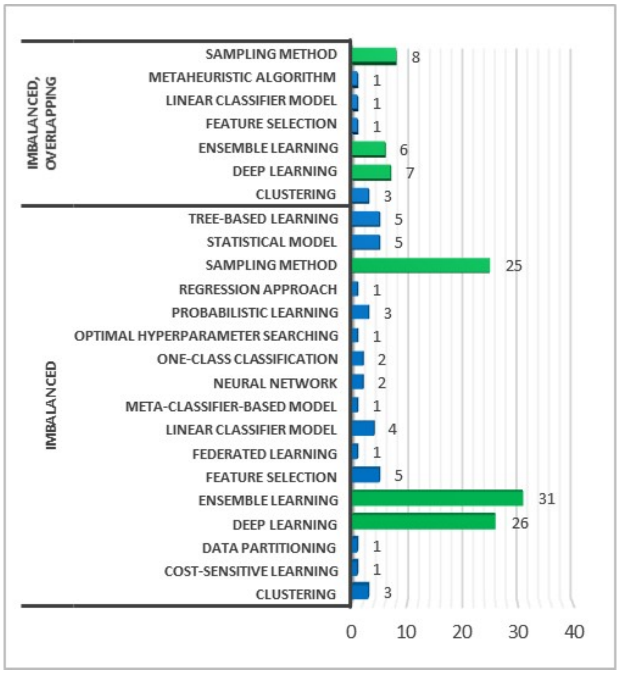

FIGURE 4. Frequency of techniques used for handling the two problems.

of papers we reviewed. Fig. 4 shows that the techniques frequently used to address only the imbalanced class distribution problem are *ensemble learning, deep learning and sampling methods*, with the frequency of 31, 26, and 25. On the other hand, sampling methods are the most frequently used techniques for addressing imbalanced class distribution and overlapping class problems with the frequency 8, followed by *deep learning and ensemble learning*, with frequencies of 7 and 6, respectively. Table 2 shows the summarized details of techniques used in the selected papers.

TABLE 2. Summary of techniques used in the selected papers.

Techniques	Description
Clustering	It is a technique that divides the data into groups based on similarities.
Cost-Sensitive Learning	It is learning that takes misclassification costs into account when training the model.
Data Partitioning	A technique to divide a dataset into test and training sets.
Deep Learning	It is a learning technique that mimics the human brain. It consists of many layers of representation and can learn, analyze and interpret a large amount of data.
Ensemble Learning	Ensemble Learning involves combining multiple learning algorithms.
Feature Selection	It is a technique to reduce irrelevant attributes in data for improving classification results.
Federated Learning	It is a technique that trains classification models using decentralized data.
Linear Classifier Model	It is a model that can classify linearly separable data.
Meta-Classifier Based Model	A model that produces a final prediction using predictions of other classifiers as inputs.
Metaheuristic Algorithm	It is a method used to solve more complex optimization problems by finding a global optimum rather than a local.
One-Class Classification	It is a technique to fit "normal" data and identify whether the new data is an anomaly.
Optimal Hyperparamete r Searching	It is a search of a model's optimal hyperparameters for optimization.
Neural Network	A shallow network that mimics the human brain to identify the relationship in a set of data.
Probabilistic Learning	It is a technique that predicts using the probability approach.
Regression Approach	It is a technique that predicts continuous values as outcomes.
Sampling Method	It is used to increase or reduce the number of data samples.
Statistical Model	It is a mathematical model that produces sample data based on statistical assumptions.
Tree-Based Learning	A technique classifies data using a tree-like structure with a series of prediction rules.

D. EVALUATION METRICS

All papers used multiple evaluation metrics. Fig. 5 demonstrates the frequency of different evaluation metrics used in the selected papers. 56 papers used True Positive Rate (TPR) (Recall or Sensitivity) as the evaluation metrics to measure the classifier's performance. 51 papers used Area Under Curve (AUC), 45 used Accuracy (Acc), and 38 used Precision (PR) and F1 as evaluation metrics. The summarized details of the evaluation metrics used by the researchers are shown in Table 3.

E. DETECTION RESULTS

We reviewed 87 papers, and the details are provided in Appendix A. We then pick the top five research works from the appendix for each of the two popular datasets and describe them in the following subsection.

F. THE EFFECTIVE MACHINE LEARNING TECHNIQUES

Appendix B summarizes the top five research works that used the two widely used datasets: European and Taiwan datasets.

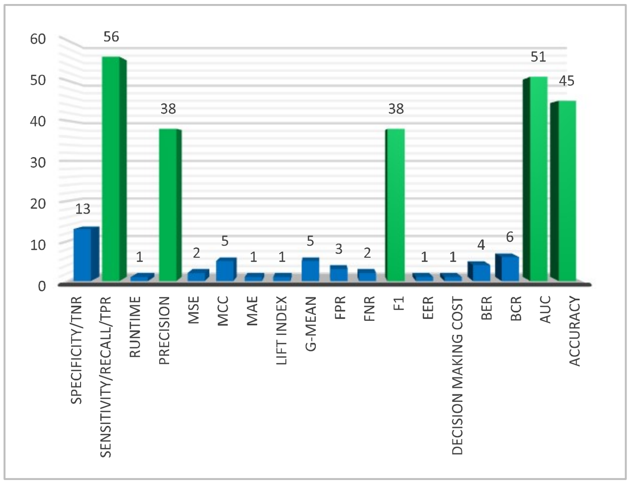

FIGURE 5. Frequency of the evaluation metrics used in the selected papers.

Our analysis shows that the most frequently used evaluation metric is the True Positive Rate (TPR/Recall/Sensitivity) (Fig. 5). Hence, we shall use this metric to compare the researchers' works and identify effective machine learning techniques for solving the two dataset problems: imbalanced class distribution and overlapping class.

Research work that achieved the best fraud detection result for the European dataset is [93]. The researchers obtained the highest result for all evaluation measurements using deep learning and neural networks. Researchers [92] and [88] used sampling methods with ensemble learning and deep learning, respectively, and both achieved a 100% True Positive Rate (TPR). Researchers [23] and [63] are not far behind, who used neural networks and achieved TPRs of 99.98% and 99.6%, respectively.

As for the Taiwan dataset, [92] obtained the highest TPR of 98.09% using sampling methods and ensemble learning for detecting payment defaults. It is followed by [88], where the researchers used sampling methods and deep learning to achieve a TPR of 93%. Researchers [20] and [63] were using sampling methods and ensemble learning. Reference [63] implemented a neural network and obtained a high TPR of 92.4% compared to [20], who achieved a TPR of 89.9%. In [61], the researchers used a linear classifier model with a sampling method and obtained a TPR of 83.6%.

V. DISCUSSION

We discuss the answers to the questions set in Section II. The widely used credit card fraud dataset is the European dataset. As for the credit card payment default, the Taiwan dataset is the most widely used. These datasets are popular because they are publicly available and widely used by many researchers.

From our findings, the imbalanced class distribution problem is widely studied by researchers, and the overlapping class problem is not as explicit as the imbalanced class distribution problem. An imbalanced class distribution can easily be identified through the number of samples in each class. On the other hand, the overlapping class problem requires a

TABLE 3. Summary of evaluation metrics used in the selected papers.

Evaluation Metrics	Description
Accuracy (Acc)	– A widely used metric – Measures the overall classification rate – Suitable for balanced datasets – Not suitable for imbalanced datasets as it may be biased towards majority classes
Precision (PR)	– Measures the number of predicted positive samples that belong to positive classes
True Positive Rate (TPR) /Recall/Sensitivity	– Measures the proportion of actual positive class samples, i.e., frauds and payment default, that are predicted as positive class
True Negative Rate (TNR)/Specificity	– Measures the proportion of actual negative class samples, i.e., normal/benign, that are predicted as negative class
Area Under the Curve (AUC)	– Measure the capability of a classifier in classifying data samples correctly.
Balanced Classification Rate (BCR)	– It calculates by taking the addition of specificity and sensitivity divided by two – balanced out Accuracy in imbalanced datasets.
Balanced Error Rate (BER)	– Calculates the average errors in each class
Equal Error Rate (EER)	– It measures by finding the equal or minimum distance between false-positive and false-negative rates. – The lower the Equal Error Rate, the higher the probability of a classification model predicting correctly.
F1 Scores (F1)	– Measures the harmonic mean of precision and recall.
False Negative Rate (FNR)	– Measures the prediction accuracy of important class samples, i.e., frauds and payment defaults that are classified incorrectly.
False Positive Rate (FPR)	– Measures the prediction accuracy of the normal/benign class samples that are classified incorrectly as frauds or payment defaults.
G-Mean	– Balance the classification performance between the majority class and the minority class. – G-mean is important in avoiding the overfitting of the majority class and underfitting of the minority class.
Mean Absolute Error (MAE)	– It measures the magnitude of error by calculating the difference between the prediction and the actual value.

TABLE 3. ***(Continued)*** **Summary of evaluation metrics used in the selected papers.**

Mean Squared Error (MSE)	– It measures the magnitude of error in a statistical model.
Matthews Correlation Coefficient (MCC)	– A statistical tool for model evaluation – Produce a high score if a prediction scores well for all categories in the confusion matrix.
Decision Making Cost	– Determine the cost when involved in decision-making.
Runtime	– Measures the time taken for training or testing a classification model.
Lift Index	– Used to measure the performance of a model based on the lift curve. The greater the lift value, the better the model is.

thorough dataset analysis, and the problem complexity may vary on different datasets. The next two paragraphs discuss why researchers widely use specific machine learning techniques to tackle the two problems.

Deep learning and neural networks can handle highly imbalanced datasets. A study by [100] shows that neural networks are capable of classifying highly imbalanced data by taking into account the direction of the unit gradient for both majority and minority classes. In addition, a local boundary expansion strategy is taken into consideration to address the issue of the minority class's inadequate empirical representation. Even without considering the unit gradient direction of majority and minority classes, coupling deep learning and neural networks with sampling methods also tackles the imbalanced datasets well. Sampling methods, i.e., under-sampling and over-sampling, work by balancing a dataset's class distribution. Particularly, over-sampling the minority class will make the decision boundary of the minority class clear. On the other hand, under-sampling reduces the overwhelming effect of the majority class. As a result, general learning algorithms can learn and pick up the minority class relatively easily than before applying sampling methods on credit card datasets.

Ensemble learning, i.e., boosting, bagging, and stacking, works by training multiple machine learning algorithms and combining detection outcomes. As single learning algorithms are usually limited and make errors, ensemble learning can help improve detections by taking advantage of the strengths of different learning algorithms [101]. Further, ensemble learning can easily be coupled with sampling methods, improve detection with a new aggregating strategy, or integrate itself with another ensemble learning for hybrid learning strategies [102].

We further discuss which evaluation metrics are appropriate to use. TPR (Recall or Sensitivity), AUC, Precision and F1 are suitable for imbalanced credit card datasets. TPR is the most widely used metric because it can measure the detection performance of a machine learning technique for majority class (normal transactions) and minority class (frauds/ payment default) separately. More importantly, calculating TPR includes false negatives, which consider frauds/payment defaults wrongly detected as normal transactions. Even though Accuracy is a very common metric to use, it is not suitable as the sole evaluation metric for imbalanced datasets because it looks into the overall performance of a machine learning technique on all dataset classes. The detection performance for frauds and payment defaults is hidden in Accuracy and may be biased towards the high detection rates for normal transactions. Researchers favor machine learning techniques that give high detection rates for credit card fraud and payment defaults.

VI. CONCLUSION

Imbalanced class distribution and overlapping class samples are the two main problems of credit card datasets. General machine learning algorithms may not be able to effectively detect the important minority classes (fraud and payment default) as the algorithms are usually biased towards the majority class samples (normal transactions). In addition, the overlapping class problem further affects the performance of the machine learning algorithms.

We analyzed the two widely used credit card datasets, European and Taiwan, and identified effective machine learning techniques to tackle the two problems so that detecting frauds and payment defaults becomes effective. Based on our analysis and review, the top three common effective techniques are (i) deep learning and neural networks, (ii) ensemble learning, and (iii) sampling methods. We recommend these techniques for credit card fraud and payment default datasets and also for future research dealing with other types of credit card data that also inherited the two problems.

The most common metrics for evaluating the performance of learning algorithms are TPR (Recall or Sensitivity), followed by Accuracy, AUC, Precision and F1. However, since the credit card datasets are usually imbalanced in their class distribution, using only Accuracy as the evaluation metric is inappropriate. If Accuracy is to be used to describe the performance of a machine learning algorithm on these imbalanced datasets, it would be better to use these four evaluation metrics (TPR, AUC, Precision and F1 scores) to complement Accuracy.

To conclude, this systematic review can become a reference for the datasets, machine learning techniques, and evaluation metrics related to future credit card research dealing with imbalanced class distribution and overlapping class problems.

APPENDIX A

See Table 4.

APPENDIX B

See Table 5.

TABLE 4. The list of 87 articles reviewed in this study.

No.	Ref.	Dataset Type	Problem(s)	Technique(s)	Technique Categories	Evaluation Metrics	Result
1	[9]	CCF	Imbalanced	• Deep Neural Network	• Deep Learning	AUC, Accuracy	AUC: 90.1 Accuracy: 95.7
2	[10]	CCF	Imbalanced	• C5.0 • Support Vector Machine (SVM) • Artificial Neural Network (ANN)	• Tree-Based Learning • Linear Classifier Model • Deep Learning	AUC, Accuracy, Sensitivity	C5.0: 60 SVM: 63 ANN: 62
3	[11]	CCF	Imbalanced	• Optimized light gradient boosting machine (OLightGBM)	• Tree-Based Learning	AUC, Accuracy, Recall, Precision, F1	AUC: 90.94 Accuracy: 98.40 Recall: 40.59 Precision: 97.34 F1: 56.95
4	[12]	CCF	Imbalanced	• Cost Sensitive Neural Network (CSNN)	• Deep Learning • Metacost - Ensemble Learning	TPR	61.40
5	[13]	CCF	Imbalanced	• Multilayer Perceptron (MLP) • SMOTE	• Deep Learning • Sampling Method	Recall, Accuracy, Precision	Accuracy: 97.84 Recall: 96.35 Precision: 99.32
6	[14]	CCF	Imbalanced, Overlapping	• Whale algorithm optimized BP neural network	• Deep Learning	Accuracy Recall Precision F-score	Accuracy: 96.4 Recall: 98.25 Precision: 97.83 F-Score: 98.04
7	[15]	CCF	Imbalanced	• AdaBoost • Majority voting	• Ensemble Learning	MCC	100
8	[16]	CCF	Imbalanced	• Bayesian Network	• Probabilistic Graphical Model	F-measure	75.33
9	[17]	CCF	Imbalanced	• Decision Tree (DT)	• Tree-Based Learning	Sensitivity	79.21
10	[18]	CCF	Imbalanced	• Under-sampling • Stacking Classifier (LR as meta-classifier)	• Sampling Method • Ensemble Learning	TPR	94
11	[19]	CCF	Imbalanced	• Oversampling • Logistic Regression (LOR)	• Sampling Method • Statistical Model	Precision, Recall, F1	Precision: 96 Recall:96 F1: 96
12	[20]	CCDP	Imbalanced	• Gradient Boosted Decision Tree (GBDT) • K-means SMOTE oversampling method	• Sampling Method • Ensemble Learning	Accuracy, Precision, Recall, F1, ROC, G-mean	Accuracy: 87.6 Precision: 85.7 Recall: 89.9 F1: 87.8 ROC: 0.93 G-mean: 87.8
13	[21]	CCF	Imbalanced	• SVM SMOTE (Oversampling) • Random Forest (RF)	• Sampling Method • Ensemble Learning	Recall, Precision, F1	Recall: 80 Precision: 91 F1: 85
14	[22]	CCF	Imbalanced	• Non-Overlapped Risk based Bagging Ensemble (NRBE)	• NRBE - Ensemble Learning	BCR, BER, Recall	BCR: 70 BER: 30 Recall: 84
15	[23]	CCDP	Imbalanced	• Distributed Neural Network Model (DDNN)	• Neural Network	Accuracy, Precision, Recall, F1	Accuracy: 99.94 Precision: 99.96 Recall: 99.98 F1: 99.97

TABLE 4. *(Continued.)* **The list of 87 articles reviewed in this study.**

16	[24]	CCF	Imbalanced	• Oversampling • Logistic Regression (LOR)	• Sampling Method • Statistical Model	AUC, Precision	AUC: 91.07 Precision: 84.13
17	[25]	CCF	Imbalanced	• Gaussian Mixture Undersampling (GMUS) • Bagging	• Sampling Method • Ensemble Learning	AUC	83.75
18	[26]	CCF	Imbalanced	• Bagging Multiple Boosted Trees (BMBT)	• Ensemble Learning	Recall, AUC, BCR, BER, MCC, TPR, TNR	Recall: 77 AUC: 85 BCR: 84 BER: 15 MCC: 69 TPR: 77 TNR: 92
19	[27]	CCF	Imbalanced	• Optimized XGBoost (OXGBoost)	• Ensemble Learning	AUC , G-mean, BCR, F1	IEEE-CIS AUC: 96.6 F1: 81.6 G-mean: 84.44 BCR: 85.69 European: AUC: 98.1 F1: 87.3 G-mean: 93.44 BCR: 93.35
20	[28]	CCF	Imbalanced	• DeepBalance	• Deep Learning	TPR, TNR, AUC	TPR: 81.76 TNR: 99.46 AUC: 97.76
21	[29]	CCF	Imbalanced	• SMOTE ENN • Bagging	• Sampling Method (Hybrid) • Ensemble Learning	AUC, Sensitivity, Specificity, G-mean	AUC: 94.98 Sensitivity: 91.10 Specificity: 99.14 G-mean: 95.02
22	[30]	CCF	Imbalanced	• Data cleaning • Feature selection • Data normalization • Data parsing • Real-time tree based metaclassifier (TBMC) • Random Forest based Level 1 Prediction • Ensemble based (Decision Tree & Gradient Boosted Tree) Level 2 Prediction	• Meta-Classifier-Based Model	FPR, FNR, Recall, AUC, BCR, BER, MCC, TPR, TNR	FPR: 8 FNR: 19 Recall: 81 AUC: 86 BCR: 86 BER: 13 MCC: 73 TPR: 81 TNR: 92
23	[31]	CCDP	Imbalanced	• Cost-sensitive AdaBoost	• Cost-Sensitive Learning • Ensemble Learning	AUC	75.61
24	[32]	CCF	Imbalanced	• Ensemble learning framework based on training set partitioning and clustering.	• Data Partitioning • Clustering • Ensemble Learning	AUC	96.52
25	[33]	CCF	Imbalanced, Overlapping	• Random forest • Random undersampling • Auto-Encoder • Artificial Neural Network	• Hybrid Ensemble Method	F1, AUC_ PR	F1:73 AUC _PR: 63
26	[34]	CCF	Imbalanced	• Undersampling • Logistic Regression (LOR)	• Sampling Method • Ensemble Learning	Precision, Recall	Precision: 95.7 Recall: 93.87
27	[35]	CCDP	Imbalanced	• RUS and MRN (MLP, Radial Basis Function-RBF, NB apply on AdaBoost.M1)	• Sampling Method • Ensemble Learning	Sensitivity, Specificity, Accuracy	Sensitivity: 53.36 Specificity: 83.1 Accuracy: 79.73

TABLE 4. ***(Continued.)*** **The list of 87 articles reviewed in this study.**

28	[36]	CCF	Imbalanced	• SMOTE • Logistic Regression (LOR)	• Sampling Method • Statistical Model	Accuracy, Precision, Recall, AUC	Accuracy: 97.04 Precision: 99.99 Recall: 97.04 AUC: 97.04
29	[37]	CCF	Imbalanced	• Undersampling • Lass-logistic • XGBoost	• Sampling Method • Regression Approach • Ensemble Learning	Accuracy	96.96
30	[38]	CCF	Imbalanced	• Cost-Sensitive Meta-learning Ensemble framework	• Meta-Learning Ensemble	AUC	99
31	[39]	CCF	Imbalanced	• Undersampling • Feed-forward Neural Network (FFNN)	• Sampling Method • Deep Learning	Recall	72.87
32	[40]	CCF	Imbalanced	• Cost-sensitive Risk Induced Bayesian Inference Bagging Model (RIBIB)	• Ensemble Learning	FPR, TPR, TNR, FNR, Recall, AUC, BCR, BER, Detection Rate, MCC	Brazilian: FPR: 11.9 TPR: 56.6 TNR: 88.1 FNR: 43.4 Recall: 56.6 Accuracy: 86.9 AUC: 72.4 Detection Rate: 56.6 BCR: 72.4 UCSD-FICO: FPR: 8 TPR: 100 TNR: 99.0 FNR: 0 Recall: 100 AUC: 99.0 BCR: 99.0 BER: 1.0 MCC: 85.0
33	[41]	CCF	Imbalanced	• SMOTE • One-class SVM	• Sampling Method • Linear Classifier Model	Accuracy, Precision	Accuracy: 70.09 Precision: 70.15
34	[42]	CCF	Imbalanced	• Pipelining	• Feature Selection	Accuracy, Precision, Recall, F1	Accuracy: 99.99 Precision: 84 Recall: 86 F1: 85
35	[43]	CCF	Imbalanced	• Long Short Term Memory (LSTM) • Fast Forward Neural Network (FFNN)	• Ensemble Learning • Deep Learning	Precision, Recall, F1, AUC-ROC, AUC-PR	European: Precision: 96.39 Recall: 66.51 F1: 78.66 AUC-ROC: 83.25 AUC-PR: 64.14 Brazilian: Precision: 83.69 Recall: 61.25 F1: 70.06 AUC-ROC: 80.37 AUC-PR: 51.62
36	[44]	CCF	Imbalanced	• SMOTE • SVM-Recursive Feature Elimination • GridSearchCV for Hyper-Parameters Optimization • Random Forest	• Sampling Method • Feature Selection • Optimal Hyperparameter Searching Technique • Tree-Based Learning • Ensemble Learning	Accuracy, Sensitivity, AUC-PR	Accuracy: 99 Sensitivity: 95 AUC-PR: 81

TABLE 4. *(Continued.)* **The list of 87 articles reviewed in this study.**

37	[45]	CCF	Imbalanced	• Homogeneity-Oriented Behaviour Analysis (HOBA) • Deep Belief Networks (DBN) • Convolutional Neural Network (CNN) • Recurrent Neural Network (RNN)	• Deep Learning	FPR, F1, Precision, Recall, Accuracy	F1: 56.8 Precision: 62.6 Recall: 51.96 Accuracy: 98.25 AUC: 97.6
38	[46]	CCF	Imbalanced	• Resampling • Random Forest	• Sampling Method • Ensemble Learning	Precision, Recall, F1, Accuracy	Precision: 97.59 Recall: 94.95 F1: 95.21 Accuracy:96.57
39	[47]	CCF	Imbalanced	• Generative Adversarial Networks (GANs)	• Deep Learning	Sensitivity, Specificity, Precision, F1, Accuracy	Precision: 97.87 F1: 81.78 Accuracy: 99.96 Sensitivity: 70.23 Specificity: 99.99
40	[48]	CCDP	Imbalanced	• Imbalanced Generative Adversarial Fusion Network (IGAFN)	• Deep Learning	AUC, accuracy, F1	Accuracy: 85.65 AUC: 73.57 F1: 61.01
41	[49]	CCF	Imbalanced, Overlapping	• Transaction Window Bagging (TWB) model	• Ensemble Learning	FPR, TPR, TNR, FNR, Recall, AUC, BCR, Fraud Detection Rate, Cost, Precision	Brazilian: FPR: 14.55 TPR: 69.32 TNR: 85.44 FNR: 30.67 Recall: 69.32 AUC: 0.7738 Fraud Detection Rate: 69.32 BCR: 77.38 Cost: 11.45 UCSD-FICO: FPR: 3 TPR: 47 TNR: 99 FNR: 52 Precision: 95.3
42	[50]	CCF	Imbalanced	• SMOTE • Deep Neural Network	• Sampling Method • Deep Learning	Accuracy, F1, Precision, Recall	Accuracy: 95.3 F1: 95.4 Precision: 95.2 Recall: 95.5 AUC: 99
43	[51]	CCF	Imbalanced	• SMOTE • CatBoost (CB)	• Sampling Method • Ensemble Learning	F1, AUC, Savings	F1: 100 AUC: 99.99 Savings: 97.10
44	[52]	CCF	Imbalanced	• RF	• Ensemble Learning	Accuracy, Precision, Recall, F1, ROC-AUC	Accuracy: 96.77 Precision: 100 Recall: 91.11 F1: 95.34 ROC-AUC: 95.55
45	[53]	CCF	Imbalanced	• Deep Learning Autoencoders • Multivariate Gaussian Distribution	• Deep Learning, Probabilistic Learning • Neural Network	Equal Error Rate (ERR)	Autoencoder: 9.8 Gaussian: 8.5
46	[54]	CCDP	Imbalanced	• Difference in Maximum Entropy (DME)	• Probabilistic Classification Model	Accuracy, F1, AUC	Accuracy: ±63 F1: ±77 AUC: 71
47	[55]	CCF	Imbalanced	• Oversampling(SMOTE)+Convolutional Neural Network(CNN)	• Sampling Method • Deep Learning	Accuracy, Precision, Recall	Accuracy: 98.9 Precision: 97 Recall: 91
48	[56]	CCF	Imbalanced	• SMOTE • Federated learning for Fraud Detection (FFD)	• Sampling Method • Federated Learning	Recall, F1, AUC, Accuracy	Recall: 78 F1: 87 AUC: 89 Accuracy: 99
49	[57]	1) CCDP 2) CCF	Imbalanced	• EasyEsamble-Boderline SMOTE-XGBoost (EEBS-XGBoost)	• Hybrid - Sampling Method	Accuracy, Recall, Precision, F1,	Taiwan: Accuracy: 69.20 Recall: 68.05 Precision: 40.18

TABLE 4. ***(Continued.)*** **The list of 87 articles reviewed in this study.**

					• Ensemble Learning	AUC, G-mean, Runtime	F1: 50.98 AUC: 76.83 G-mean: 70.03 Runtime: 4.24 European: Accuracy: 92.08 Recall: 94.83 Precision: 2.05 F1: 4.02 AUC: 98.21 G-mean: 93.44 Runtime: 5.9228
50	[58]	CCF	Imbalanced	• Random over sampling • BiLSTM-MaxPooling • BiGRU-MaxPooling	• Sampling Method • Deep Learning	AUC, Precision, Recall, F1	AUC: 91.37 Precision: 91.14 Recall: 94.59 F1: 92.81
51	[59]	CCF	Imbalanced, Overlapping	• Random Forest • XGBoost • Neural Network • Oversampling • Undersampling	• Ensemble Learning • Sampling Method	AUC_ PR	50
52	[60]	CCF	Imbalanced, Overlapping	• K-Means	• Ensemble Clustering	AUC_PR, AUC_ROC	AUC-PR: 26.56 AUC-ROC: 89.37
53	[61]	CCDP	Imbalanced	• DeepInsight , CNN, ADASYN	• DeepInsight, Deep Learning, Sampling method	Accuracy, TPR, TNR, MCC	Accuracy: 99.959 TPR: 84.158 TNR: 99.988 MCC: 0.8816
54	[62]	1) CCDP 2) CCF	Imbalanced, Overlapping	• SVM • Reduce Overlapping with ADASYN	• Sampling Method • Ensemble Learning • Linear Classifier Model.	Recall, Precision, F-measure, Gmean	Taiwan: ROA Recall: 83.6 Precision: 85.4 F1: 84 G-mean: 89.9 UCSD-FICO: ROA Recall: 94.4 Precision: 86 F1: 87.2 G-mean: 89
55	[63]	1) CCF 2) CCDP	Imbalanced, Overlapping	• LSTM Ensemble (AdaBoost) • SMOTE-ENN	• Neural Network • Ensemble Learning • Sampling	Sensitivity, Specificity, AUC	European Sensitivity: 99.6 Specificity: 99.8 AUC: 99 Taiwan Sensitivity: 92.4 Specificity: 95.1 AUC: 93
56	[64]	CCDP	Imbalanced	• Random Forest	• Ensemble Learning	Accuracy	99.27
57	[65]	CCF	Imbalanced, Overlapping	• XGBoost • SMOTE • ADASYN	• Ensemble Learning • Sampling Method	Accuracy	XGBoost+SMOTE ADASYN: 99
58	[66]	CCDP	Imbalanced	• SVM	• Linear Classifier Model	Accuracy, Precision	Taiwan Accuracy: 94.10 Precision: 95.28 German Accuracy: 88.74 Precision: 89.92
59	[67]	CCF	Imbalanced	• Autoencoder	• Neural Network	AUC, TPR	AUC: 88.24 TPR: 90.26
60	[68]	CCF	Imbalanced	• GAN	• Deep Learning	Precision, Recall, Accuracy, F1	Credit card dataset Precision: 90.97 Recall: 91.74 Accuracy: 91.30 F1: 91.33

TABLE 4. *(Continued.)* **The list of 87 articles reviewed in this study.**

							European Precision: 91.89 Recall: 85.00 Accuracy: 97.95 F1: 88.31 Online payment dataset Precision: 97.73 Recall: 87.18 Accuracy: 98.68 F1: 92.15
61	[69]	CCF	Imbalanced	• KNN-DL	• Ensemble Learning	Accuracy, Precision, Recall	Accuracy: 99.59 Precision: 99.59 Recall: 99.59
62	[70]	CCF	Imbalanced	• Particle Swarm Optimization (PSO) • Auto-Associative Neural Network (AANN)	• One-Class Classification	Classification Rate	TPR: 95.31
63	[71]	CCF	Imbalanced	• One-class Adversarial Fraud Detection Nets (CS-OCAN) • Autoencoders and Complementary generative adversarial networks (GAN)	• One-Class Classification • Deep Learning	Recall, Precision, Accuracy, F1	Precision: 93.7 Recall: 85.6 Accuracy: 90.9 F1: 89.4
64	[72]	CCF	Imbalanced	• Cross-Silo Federated XGBoost	• Ensemble Learning	Accuracy, AUC, F1	Accuracy: 99.97 AUC: 97.89 F1: 90.14
65	[73]	CCF	Imbalanced	• Competitive Swarn Optimization - Deep Convolutional Neural Network	• Feature Selection • Artificial Neural Network	Accuracy, Absolute Mean Error (MAE), Mean Squared Error (MSE)	Accuracy: 98.2 MAE: 13.1 MSE: 26.3
66	[74]	CCF	Imbalanced	• Genetic Algorithm • Random Forest	• Feature Selection • Ensemble Learning	Accuracy, Recall, Precision, F1, AUC	Accuracy: 99.95 Recall: 75.75 Precision: 87.48 F1: 86.23 AUC: 0.96
67	[75]	CCF	Imbalanced	• Convolutional Neural Network (CNN)	• Deep Learning	Accuracy	99.72
68	[76]	CCF	Imbalanced	• Logistic Regression	• Statistical Model	AUC	95.55
69	[77]	CCF	Imbalanced	• Logistic Regression	• Statistical Model	Accuracy, Precision, Recall, F1	Accuracy: 99.05 Precision: 81.94 Recall: 52.21 F1: 67
70	[78]	CCF	Imbalanced	• PSO weighted K-means clustering hybrid, SMOTENN	• Clustering, Sampling method	Accuracy, F1	Accuracy: 73.17 F1: 63.33
71	[79]	CCF	Imbalanced	• Artificial Neural Network	• Neural Network	Accuracy, Precision, Recall	Accuracy: 99.92 Precision: 81.15 Recall: 76.19
72	[80]	CCF	Imbalanced	• Autoencoder	• Neural Network	F1, Recall, Precision, AUC-ROC, AUC-PR	European: F1: 82.63 Recall: 71.31 Precision: 94.56 AUC-ROC: 90.93 AUC-PR: 77.46 German: F1: 64 Recall: 72.73 Precision: 57.14 AUC-ROC: 74.67 AUC-PR: 52.28
73	[81]	CCDP	Imbalanced	• Recurrent Neural Network (RNN), GRU	• Neural Network, Deep Learning	Lift Index, AUC, Accuracy	Lift Index: 65.9 AUC: 78.2 Accuracy: 80.2

TABLE 4. *(Continued.)* **The list of 87 articles reviewed in this study.**

74	[82]	CCF	Imbalanced	• Random Forest • RUS • Convolutional Autoencoder (CAE)	• Ensemble Learning • Sampling Method • Deep Learning	F1, AUC	F1: 90.2 AUC: 98.4
75	[83]	CCF	Imbalanced, Overlapping	• Spiral oversampling balancing technique (SOBT)	• Sampling Method	AUC	European-AUC: 90.47 IEEE-CIS-AUC: 72.10
76	[84]	CCF	Imbalanced, Overlapping	• Undersampling • Fuzzy C-means • Similarity-Based Selection (SBS)	• Sampling Method • Clustering	Accuracy, Precision, F1, Sensitivity, Specificity, AUC	SBS-ANN Accuracy: 96.6 Precision: 96.4 F1: 93 Sensitivity: 89.9 Specificity: 98.9 AUC: 94.4
77	[85]	CCF	Imbalanced, Overlapping	• Support Vector Data Description (SVDD) with hyperspace optimization using Polynomial Self Learning PSO and feature selection	• Feature Selection • Metaheuristic Algorithm	Accuracy, Precision, Recall, F1, AUC	Accuracy: 93 Precision: 90 Recall: 97 F1: 93 AUC: 91.3
78	[86]	CCF	Imbalanced	• Multilayer Perceptron (MLP)	• Neural Network	Mean Squared Error (MSE)	MSE: 0.18
79	[87]	CCF	Imbalanced	• Extreme Gradient Boosting • Random Under-Sampling	• Ensemble learning • Sampling Method	AUC	AUC: 99
80	[88]	1) CCF 2) CCDP	Imbalanced, Overlapping	• LSTM • GRU • MLP • SMOTE-ENN	• Deep Learning, • Sampling Method	Sensitivity, Specificity, AUC	European Sensitivity: 100 Specificity: 99.7 AUC: 100 Taiwan Sensitivity: 93 Specificity: 96.1 AUC: 39.40
81	[89]	CCF	Imbalanced	• Support Vector Machines	• Linear Classifier Model	Precision, Recall, F1	Precision: 98 Recall: 95 F1: 96
82	[90]	CCF	Imbalanced, Overlapping	• Autoencoders • Deep Convolutional Neural Network (ADASYN-VAE-CNN)	• Neural Network • Deep Learning	Recall, AUC	Recall: 94.11 AUC: 97.3
83	[91]	CCF	Imbalanced	• Random Forest	• Ensemble Learning	Accuracy, F1, Recall, Precision, Specificity	Accuracy: 96 F1: 17 Recall: 97 Precision: 9 Specificity: 96
84	[92]	CCF, CCDP	Imbalanced, Overlapping	• XGBoost, stratified sampling	• Meta-learning ensemble, sampling method	Accuracy, Precision, Recall, F1, AUC	**European** Accuracy: 99.98 Precision: 99.96 Recall: 100 F1: 99.95 AUC: 98.94 **Taiwan** Accuracy: 99 Precision: 99.52 Recall: 98.09 F1: 98.83 AUC: 99
85	[93]	CCF	Imbalanced, Overlapping	• Long Short-Term Memory-Recurrent Neural Network (LSTM-RNN)	• Deep Learning • Neural Network	Precision, Recall, F1, Sensitivity, Accuracy	Precision: 100 Recall: 100 F1: 100 Sensitivity:100 Accuracy: 100
86	[94]	CCF	Imbalanced	• Weighted Average	• Ensemble Learning • Feature Selection	AUC, Recall, F1, Precision, Accuracy	AUC: 95.2 Recall: 90.65 F1: 91.67 Precision: 92.79 Accuracy: 99.44
87	[95]	CCF	Imbalanced	• Deep Boosting Decision Trees (DBDT),	• Deep Learning • Tree-Based Learning	AUC	99.83

TABLE 5. The high detection rates achieved by the 321 researchers' work using the popular credit 322 card datasets. the numbers in bold show the 323 highest tpr achieved by research work utilizing 324 a particular dataset.

Dataset	Research	Year	Addressed Issue(s)	Technique(s)	Technique Categories	TPR % (F/PD)*	Acc. %	AUC %	PR %	F1 %
European (Fraud)	[93]	2022	• Imbalanced • Overlapping	• Long Short-Term Memory-Recurrent Neural Network (LSTM-RNN)	• Neural Network • Deep Learning	**100**	100	-	100	100
	[92]	2023	• Imbalanced • Overlapping	• XGBoost • Stratified Sampling	• Meta-learning ensemble • Sampling Method	**100**	99.98	98.94	99.96	99.95
	[88]	2023	• Imbalanced • Overlapping	• LSTM • GRU • MLP • SMOTE-ENN	• Deep Learning • Sampling Method	**100**	-	100	-	-
	[23]	2023	• Imbalanced	• Distributed Neural Network Model (DDNN)	• Neural Network	99.98	99.94	-	99.96	99.97
	[63]	2022	• Imbalanced • Overlapping	• LSTM Ensemble (AdaBoost) • SMOTE-ENN	• Neural Network • Ensemble Learning • Sampling	99.60	-	99	-	-
Taiwan (Payment Default)	[92]	2023	• Imbalanced • Overlapping	• XGBoost • Stratified Sampling	• Meta-learning ensemble • Sampling Method	**98.09**	99.00	99.00	99.52	98.83
	[88]	2023	• Imbalanced • Overlapping	• LSTM • GRU • MLP • SMOTE-ENN	• Deep Learning • Sampling Method	93	-	39.4	-	-
	[63]	2022	• Imbalanced • Overlapping	• LSTM Ensemble (AdaBoost) • SMOTE-ENN	• Neural Network • Ensemble Learning • Sampling Method	92.4	-	93	-	-
	[20]	2020	• Imbalanced	• SMOTE • Gradient Boosted Decision Tree	• Sampling Method • Ensemble Learning	89.9	87.6	-	85.7	87.8
	[61]	2021	• Imbalanced • Overlapping	• SVM, Reduce Overlapping with ADASYN	• Linear classifier model, Sampling Method	83.60	-	-	85.40	84.00

**F/PD – Fraud/Payment Default*

REFERENCES

[1] N. Lee. (2021). *Credit Card Fraud Will Increase Due To the Covid Pandemic, Experts Warn*. CNBC. [Online]. Available: https://www.cnbc.com/2021/01/27/credit-card-fraud-is-on-the-rise-due-to-covid-pandemic.html

[2] M. Shanmugam. (2021). *Alternative Views: Cybercrime on the Rise Amid the Pandemic*. [Online]. Available: https://www.theedgemarkets.com/article/alternative-views-cybercrime-rise-amid-pandemic

[3] (2022). *Nilson Newsletter*. [Online]. Available: https://nilsonreport.com/newsletters/1232/

[4] (2022). *BNM Annual Report 2022*. [Online]. Available: https://www.bnm.gov.my/documents/20124/10150308/ar2022_en_book.pdf

[5] Chester Tay. (2022). *Malaysians Use Credit Cards More and Borrow More*. Edge Malaysia. [Online]. Available: https://theedgemalaysia.com/article/malaysians-use-credit-cards-more-and-borrow-more

[6] W. Wu. (2018). *What Happens If You Do Not Pay Back Your Credit Card Debt?*. [Online]. Available: https://ringgitplus.com/en/blog/credit-cards/what-happens-if-you-do-not-pay-back-your-credit-card-debt.html

[7] D. Moher, P.-P. Group, L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement," *Systematic Rev.*, vol. 4, no. 1, pp. 1–9, Dec. 2015.

[8] C. M. D. C. Santos, C. A. D. M. Pimenta, and M. R. C. Nobre, "The PICO strategy for the research question construction and evidence search," *Revista Latino-Americana de Enfermagem*, vol. 15, no. 3, pp. 508–511, Jun. 2007, doi: 10.1590/s0104-11692007000300023.

[9] X. Yu, X. Li, Y. Dong, and R. Zheng, "A deep neural network algorithm for detecting credit card fraud," in *Proc. Int. Conf. Big Data, Artif. Intell. Internet Things Eng. (ICBAIE)*, Jun. 2020, pp. 181–183, doi: 10.1109/ICBAIE49996.2020.00045.

[10] S. Makki, Z. Assaghir, Y. Taher, R. Haque, and H. Zeineddine, "An experimental study with imbalanced classification approaches for credit card fraud detection," *IEEE Access*, vol. 7, pp. 93010–93022, 2019, doi: 10.1109/ACCESS.2019.2927266.

[11] 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.

[12] F. Ghobadi and M. Rohani, "Cost sensitive modeling of credit card fraud using neural network strategy," in *Proc. 2nd Int. Conf. Signal Process. Intell. Syst. (ICSPIS)*, Dec. 2016, pp. 1–5, doi: 10.1109/ICSPIS.2016.7869880.

[13] F. Z. El hlouli, J. Riffi, M. A. Mahraz, A. El Yahyaouy, and H. Tairi, "Credit card fraud detection based on multilayer perceptron and extreme learning machine architectures," in *Proc. Int. Conf. Intell. Syst. Comput. Vis. (ISCV)*, 2020, pp. 1–5, doi: 10.1109/iscv49265.2020.9204185.

[14] C. Wang, Y. Wang, Z. Ye, L. Yan, W. Cai, and S. Pan, "Credit card fraud detection based on whale algorithm optimized BP neural network," in *Proc. 13th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Aug. 2018, pp. 1–4, doi: 10.1109/ICCSE.2018.8468855.

[15] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, "Credit card fraud detection using AdaBoost and majority voting," *IEEE Access*, vol. 6, pp. 14277–14284, 2018, doi: 10.1109/ACCESS.2018.2806420.

[16] A. G. C. de Sá, A. C. M. Pereira, and G. L. Pappa, "A customized classification algorithm for credit card fraud detection," *Eng. Appl. Artif. Intell.*, vol. 72, pp. 21–29, Jun. 2018, doi: 10.1016/j.engappai.2018.03.011.

[17] S. Khatri, A. Arora, and A. P. Agrawal, "Supervised machine learning algorithms for credit card fraud detection: A comparison," in *Proc. 10th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2020, pp. 680–683, doi: 10.1109/Confluence47617.2020.9057851.

[18] S. Dhankhad, E. Mohammed, and B. Far, "Supervised machine learning algorithms for credit card fraudulent transaction detection: A comparative study," in *Proc. IEEE Int. Conf. Inf. Reuse Integr. (IRI)*, Jul. 2018, pp. 122–125, doi: 10.1109/IRI.2018.00025.

[19] A. Deshpande, C. Kamath, and M. Joglekar, "A comparison study of classification methods and effects of sampling on unbalanced data," in *Proc. Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Nov. 2019, pp. 1056–1063, doi: 10.1109/ICSSIT46314.2019.8987801.

[20] T. M. Alam, K. Shaukat, I. A. Hameed, S. Luo, M. U. Sarwar, S. Shabbir, J. Li, and M. Khushi, "An investigation of credit card default prediction in the imbalanced datasets," *IEEE Access*, vol. 8, pp. 201173–201198, 2020, doi: 10.1109/ACCESS.2020.3033784.

[21] S. Taneja, B. Suri, and C. Kothari, "Application of balancing techniques with ensemble approach for credit card fraud detection," in *Proc. Int. Conf. Comput., Power Commun. Technol. (GUCON)*, Sep. 2019, pp. 753–758.

[22] S. Akila and U. S. Reddy, "Credit card fraud detection using non-overlapped risk based bagging ensemble (NRBE)," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC)*, Dec. 2017, pp. 1–4, doi: 10.1109/ICCIC.2017.8524418.

[23] Y.-T. Lei, C.-Q. Ma, Y.-S. Ren, X.-Q. Chen, S. Narayan, and A. N. Q. Huynh, "A distributed deep neural network model for credit card fraud detection," *Finance Res. Lett.*, vol. 58, Dec. 2023, Art. no. 104547, doi: 10.1016/j.frl.2023.104547.

[24] M. Puh and L. Brkic, "Detecting credit card fraud using selected machine learning algorithms," in *Proc. 42nd Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO)*, May 2019, pp. 1250–1255, doi: 10.23919/MIPRO.2019.8757212.

[25] F. Zhang, G. Liu, Z. Li, C. Yan, and C. Jiang, "GMM-based undersampling and its application for credit card fraud detection," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8, doi: 10.1109/IJCNN.2019.8852415.

[26] M. Kavitha and M. Suriakala, "Hybrid multi-level credit card fraud detection system by bagging multiple boosted trees (BMBT)," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC)*, Dec. 2017, pp. 1–5, doi: 10.1109/ICCIC.2017.8524161.

[27] C. V. Priscilla and D. P. Prabha, "Influence of optimizing XGBoost to handle class imbalance in credit card fraud detection," in *Proc. 3rd Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Aug. 2020, pp. 1309–1315, doi: 10.1109/ICSSIT48917.2020.9214206.

[28] P. Xenopoulos, "Introducing DeepBalance: Random deep belief network ensembles to address class imbalance," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2017, pp. 3684–3689, doi: 10.1109/BIGDATA.2017.8258364.

[29] D. S. Sisodia, N. K. Reddy, and S. Bhandari, "Performance evaluation of class balancing techniques for credit card fraud detection," in *Proc. IEEE Int. Conf. Power, Control, Signals Instrum. Eng. (ICPCSI)*, Sep. 2017, pp. 2747–2752, doi: 10.1109/ICPCSI.2017.8392219.

[30] M. Kavitha and M. Suriakala, "Real time credit card fraud detection on huge imbalanced data using meta-classifiers," in *Proc. Int. Conf. Inventive Comput. Informat. (ICICI)*, Nov. 2017, pp. 881–887, doi: 10.1109/ICICI.2017.8365263.

[31] M. Lv, Y. Ren, and Y. Chen, "Research on imbalanced data: Based on SMOTE-AdaBoost algorithm," in *Proc. 3rd Int. Conf. Electron. Inf. Technol. Comput. Eng. (EITCE)*, Oct. 2019, pp. 1165–1170, doi: 10.1109/EITCE47263.2019.9094859.

[32] H. Wang, P. Zhu, X. Zou, and S. Qin, "An ensemble learning framework for credit card fraud detection based on training set partitioning and clustering," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Advanced Trusted Comput., Scalable Comput. Commun., Cloud Big Data Computing, Internet People Smart City Innov.*, 2018, pp. 94–98, doi: 10.1109/smartworld.2018.00051.

[33] Z. Li, M. Huang, G. Liu, and C. Jiang, "A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection," *Exp. Syst. Appl.*, vol. 175, Aug. 2021, Art. no. 114750, doi: 10.1016/j.eswa.2021.114750.

[34] A. Mishra and C. Ghorpade, "Credit card fraud detection on the skewed data using various classification and ensemble techniques," in *Proc. IEEE Int. Students' Conf. Electr., Electron. Comput. Sci. (SCEECS)*, Feb. 2018, pp. 1–5, doi: 10.1109/SCEECS.2018.8546939.

[35] A. Charleonnan, "Credit card fraud detection using RUS and MRN algorithms," in *Proc. Manage. Innov. Technol. Int. Conf. (MITicon)*, Oct. 2016, pp. 73–76, doi: 10.1109/MITICON.2016.8025244.

[36] J. V. Sriram Sasank, G. R. Sahith, K. Abhinav, and M. Belwal, "Credit card fraud detection using various classification and sampling techniques: A comparative study," in *Proc. Int. Conf. Commun. Electron. Syst. (ICCES)*, 2019, pp. 1713–1718, doi: 10.1109/icces45898.2019.9002289.

[37] D. Li, "Credit card fraud identification based on unbalanced data set based on fusion model," in *Proc. IEEE 1st Int. Conf. Civil Aviation Saf. Inf. Technol. (ICCASIT)*, Oct. 2019, pp. 235–239, doi: 10.1109/ICCASIT48058.2019.8973167.

[38] T. A. Olowookere and O. S. Adewale, "A framework for detecting credit card fraud with cost-sensitive meta-learning ensemble approach," *Sci. Afr.*, vol. 8, Jul. 2020, Art. no. e00464, doi: 10.1016/j.sciaf.2020.e00464.

[39] E. Kim, J. Lee, H. Shin, H. Yang, S. Cho, S.-K. Nam, Y. Song, J.-A. Yoon, and J.-I. Kim, "Champion-challenger analysis for credit card fraud detection: Hybrid ensemble and deep learning," *Exp. Syst. Appl.*, vol. 128, pp. 214–224, Aug. 2019, doi: 10.1016/j.eswa.2019.03.042.

[40] S. Akila and U. Srinivasulu Reddy, "Cost-sensitive risk induced Bayesian inference bagging (RIBIB) for credit card fraud detection," *J. Comput. Sci.*, vol. 27, pp. 247–254, Jul. 2018, doi: 10.1016/j.jocs.2018.06.009.

[41] V. N. Dornadula and S. Geetha, "Credit card fraud detection using machine learning algorithms," *Proc. Comput. Sci.*, vol. 165, pp. 631–641, Jan. 2019, doi: 10.1016/j.procs.2020.01.057.

[42] S. Bagga, A. Goyal, N. Gupta, and A. Goyal, "Credit card fraud detection using pipeling and ensemble learning," *Proc. Comput. Sci.*, vol. 173, pp. 104–112, Jan. 2020, doi: 10.1016/j.procs.2020.06.014.J.

[43] J. Forough and S. Momtazi, "Ensemble of deep sequential models for credit card fraud detection," *Appl. Soft Comput.*, vol. 99, Feb. 2021, Art. no. 106883, doi: 10.1016/j.asoc.2020.106883.

[44] N. Rtayli and N. Enneya, "Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization," *J. Inf. Secur. Appl.*, vol. 55, Dec. 2020, Art. no. 102596, doi: 10.1016/j.jisa.2020.102596.

[45] X. Zhang, Y. Han, W. Xu, and Q. Wang, "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture," *Inf. Sci.*, vol. 557, pp. 302–316, May 2021, doi: 10.1016/j.ins.2019.05.023.

[46] A Sahu, G. M. Harshvardhan, and M. K. Gourisaria, "A dual approach for credit card fraud detection using neural network and data mining techniques," in *Proc. IEEE 17th India Council Int. Conf. (INDICON)*, Dec. 2020, pp. 1–7, doi: 10.1109/INDICON49873.2020.9342462.

[47] U. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, "Using generative adversarial networks for improving classification effectiveness in credit card fraud detection," *Inf. Sci.*, vol. 479, pp. 448–455, Apr. 2019, doi: 10.1016/j.ins.2017.12.030.

[48] K. Lei, Y. Xie, S. Zhong, J. Dai, M. Yang, and Y. Shen, "Generative adversarial fusion network for class imbalance credit scoring," *Neural Comput. Appl.*, vol. 32, no. 12, pp. 8451–8462, Jun. 2020, doi: 10.1007/s00521-019-04335-1.

[49] A. Somasundaram and S. Reddy, "Parallel and incremental credit card fraud detection model to handle concept drift and data imbalance," *Neural Comput. Appl.*, vol. 31, no. S1, pp. 3–14, Jan. 2019, doi: 10.1007/s00521-018-3633-8.

[50] K. I. Alkhatib, A. I. Al-Aiad, M. H. Almahmoud, and O. N. Elayan, "Credit card fraud detection based on deep neural network approach," in *Proc. 12th Int. Conf. Inf. Commun. Syst. (ICICS)*, May 2021, pp. 153–156, doi: 10.1109/ICICS52457.2021.9464555.

[51] D. Almhaithawi, A. Jafar, and M. Aljnidi, "Example-dependent cost-sensitive credit cards fraud detection using SMOTE and Bayes minimum risk," *Social Netw. Appl. Sci.*, vol. 2, no. 9, Sep. 2020, doi: 10.1007/s42452-020-03375-w.

[52] D. Prajapati, A. Tripathi, J. Mehta, K. Jhaveri, and V. Kelkar, "Credit card fraud detection using machine learning," in *Proc. Int. Conf. Adv. Comput., Commun., Control (ICAC)*, Dec. 2021, pp. 1–6.

[53] M. Parreno-Centeno, M. A. Ali, Y. Guan, and A. van Moorsel, "Unsupervised machine learning for card payment fraud detection," in *Proc. Int. Conf. Risks Secur. Internet Syst.* (Lecture Notes in Computer Science), 2020, pp. 247–262, doi: 10.1007/978-3-030-41568-6_16.

[54] R. Saia and S. Carta, "An entropy based algorithm for credit scoring," in *Proc. Int. Conf. Res. Practical Issues Enterprise Inf. Syst.* (Lecture Notes in Computer Science), 2016, pp. 263–276, doi: 10.1007/978-3-319-49944-4_20.

[55] A. A. El Naby, E. El-Din Hemdan, and A. El-Sayed, "Deep learning approach for credit card fraud detection," in *Proc. Int. Conf. Electron. Eng. (ICEEM)*, Jul. 2021, pp. 1–5, doi: 10.1109/ICEEM52022.2021.9480639.

[56] W. Yang, Y. Zhang, K. Ye, L. Li, and C.-Z. Xu, "FFD: A federated learning based method for credit card fraud detection," in *Proc. Int. Conf. Big Data*, 2019, pp. 18–32, doi: 10.1007/978-3-030-23551-2_2.

[57] S.-X. Zhao, X. Wang, and Q. Yue, "A novel mixed sampling algorithm for imbalanced data based on XGBoost," in *Proc. China Conf. Wireless Sensor Netw.*, 2020, pp. 181–196, doi: 10.1007/978-981-33-4214-9_14.

[58] H. Najadat, O. Altiti, A. A. Aqouleh, and M. Younes, "Credit card fraud detection based on machine and deep learning," in *Proc. 11th Int. Conf. Inf. Commun. Syst. (ICICS)*, Apr. 2020, doi: 10.1109/icics49469.2020.239524.

[59] A. Alazizi, A. Habrard, F. Jacquenet, L. He-Guelton, F. Oblé, and W. Siblini, "Anomaly detection, consider your dataset first an illustration on fraud detection," in *Proc. IEEE 31st Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2019, pp. 1351–1355, doi: 10.1109/ICTAI.2019.00188.

[60] U. Porwal and S. Mukund, "Credit card fraud detection in e-commerce," in *Proc. 18th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./13th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2019, pp. 280–287.

[61] J.-R. Jiang and C.-K. Liao, "Credit card fraud detection based on deep insight and deep learning," in *Proc. Int. Conf. Consum. Electron.*, Jul. 2023, doi: 10.1109/icce-taiwan58799.2023.10226905.

[62] E. B. Fatima, B. Omar, E. M. Abdelmajid, F. Rustam, A. Mehmood, and G. S. Choi, "Minimizing the overlapping degree to improve class-imbalanced learning under sparse feature selection: Application to fraud detection," *IEEE Access*, vol. 9, pp. 28101–28110, 2021, doi: 10.1109/access.2021.3056285.

[63] E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba, and G. Obaido, "A neural network ensemble with feature engineering for improved credit card fraud detection," *IEEE Access*, vol. 10, pp. 16400–16407, 2022, doi: 10.1109/access.2022.3148298.

[64] Y. Yu, "The application of machine learning algorithms in credit card default prediction," in *Proc. Int. Conf. Comput. Data Sci. (CDS)*, Aug. 2020, pp. 212–218, doi: 10.1109/CDS49703.2020.00050.

[65] H. Zou, "Analysis of best sampling strategy in credit card fraud detection using machine learning," in *Proc. 6th Int. Conf. Intell. Inf. Technol.*, Feb. 2021, pp. 40–44, doi: 10.1145/3460179.3460186.

[66] K. Murugan, A. Felicia, B. Gomathy, P. T. Saravanakumar, S. M. Ramesh, and E. Sakthivel, "A credit card fraud identification technique using support vector machine," in *Proc. Int. Conf. Appl. Intell. Sustain. Comput. (ICAISC)*, Jun. 2023, doi: 10.1109/icaisc58445.2023.10199684.

[67] R. K. L. Kennedy, Z. Salekshahrezaee, and T. M. Khoshgoftaar, "A novel approach for unsupervised learning of highly-imbalanced data," in *Proc. IEEE 4th Int. Conf. Cognit. Mach. Intell. (CogMI)*, Dec. 2022, pp. 52–58, doi: 10.1109/CogMI56440.2022.00018.

[68] W. Zhang, S. Zhang, J. Chen, and H. Huang, "CCFD-Gan: Credit card fraud detection based on generative adversarial networks enhanced by penalty mechanism," in *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, 2023, pp. 1–8, doi: 10.1109/ijcnn54540.2023.10191810.

[69] D. Kaur, A. Saini, and D. Gupta, "Credit card fraud detection using machine learning, deep learning, and ensemble of the both," in *Proc. 7th Int. Conf. Parallel, Distrib. Grid Comput. (PDGC)*, Nov. 2022, pp. 484–489, doi: 10.1109/PDGC56933.2022.10053175.

[70] A. Alshammari, R. Alshammari, M. Altalak, K. Alshammari, and A. Alhakamy, "Credit-card fraud detection system using big data analytics," in *Proc. Int. Conf. Electr., Comput., Commun. Mechatronics Eng. (ICECCME)*, Nov. 2022, pp. 1–7, doi: 10.1109/ICECCME55909.2022.9987791.

[71] H. Peng, J. Zhao, L. Li, Y. Ren, and S. Zhao, "One-class adversarial fraud detection nets with class specific representations," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 6, pp. 3793–3803, Nov./Dec. 2023, doi: 10.1109/TNSE.2023.3273543.

[72] M. Yang, S. Liu, J. Xu, G. Tan, C. Li, and L. Song, "Achieving privacy-preserving cross-silo anomaly detection using federated XGBoost," *J. Franklin Inst.*, vol. 360, no. 9, pp. 6194–6210, Jun. 2023, doi: 10.1016/j.jfranklin.2023.04.002.

[73] T. Karthikeyan, M. Govindarajan, and V. Vijayakumar, "An effective fraud detection using competitive swarm optimization based deep neural network," *Meas., Sensors*, vol. 27, Jun. 2023, Art. no. 100793, doi: 10.1016/j.measen.2023.100793.

[74] 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, pp. 1–17, Dec. 2022, doi: 10.1186/s40537-022-00573-8.

[75] 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.

[76] A. Aditi, A. Dubey, A. Mathur, and P. Garg, "Credit card fraud detection using advanced machine learning techniques," in *Proc. 5th Int. Conf. Comput. Intell. Commun. Technol. (CCICT)*, Jul. 2022, pp. 56–60, doi: 10.1109/CCiCT56684.2022.00022.

[77] M. Devika, S. R. Kishan, L. S. Manohar, and N. Vijaya, "Credit card fraud detection using logistic regression," in *Proc. 2nd Int. Conf. Adv. Technol. Intell. Control, Environ., Comput. Commun. Eng. (ICATIECE)*, Dec. 2022, pp. 1–6, doi: 10.1109/ICATIECE56365.2022.10046976.

[78] N. Sharma and V. Ranjan, "Credit card fraud detection: A hybrid of PSO and K-means clustering unsupervised approach," in *Proc. 13th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2023, pp. 445–450, doi: 10.1109/Confluence56041.2023.10048876.

[79] A. Rb and S. K. Kr, "Credit card fraud detection using artificial neural network," *Global Transitions Proc.*, vol. 2, no. 1, pp. 35–41, Jun. 2021, doi: 10.1016/j.gltp.2021.01.006.

[80] H. Fanai and H. Abbasimehr, "A novel combined approach based on deep autoencoder and deep classifiers for credit card fraud detection," *Exp. Syst. Appl.*, vol. 217, May 2023, Art. no. 119562, doi: 10.1016/j.eswa.2023.119562.

[81] T.-C. Hsu, S.-T. Liou, Y.-P. Wang, and Y.-S. Huang, "Enhanced recurrent neural network for combining static and dynamic features for credit card default prediction," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1572–1576, doi: 10.1109/ICASSP.2019.8682212.

[82] Z. Salekshahrezaee, J. L. Leevy, and T. M. Khoshgoftaar, "The effect of feature extraction and data sampling on credit card fraud detection," *J. Big Data*, vol. 10, no. 1, Jan. 2023, doi: 10.1186/s40537-023-00684-w.

[83] L. Ni, J. Li, H. Xu, X. Wang, and J. Zhang, "Fraud feature boosting mechanism and spiral oversampling balancing technique for credit card fraud detection," *IEEE Trans. Computat. Social Syst.*, early access, Feb. 13, 2023, doi: 10.1109/TCSS.2023.3242149.

[84] H. Ahmad, B. Kasasbeh, B. Aldabaybah, and E. Rawashdeh, "Class balancing framework for credit card fraud detection based on clustering and similarity-based selection (SBS)," *Int. J. Inf. Technol.*, vol. 15, no. 1, pp. 325–333, Jun. 2022, doi: 10.1007/s41870-022-00987-w.

[85] A. Mniai, M. Tarik, and K. Jebari, "A novel framework for credit card fraud detection," *IEEE Access*, vol. 11, pp. 112776–112786, 2023, doi: 10.1109/ACCESS.2023.3323842.

[86] D. Dighe, S. Patil, and S. Kokate, "Detection of credit card fraud transactions using machine learning algorithms and neural networks: A comparative study," in *Proc. 4th Int. Conf. Comput. Commun. Control Autom. (ICCUBEA)*, Aug. 2018, pp. 1–6, doi: 10.1109/ICCUBEA.2018.8697799.

[87] D. S. Nijwala, S. Maurya, M. P. Thapliyal, and R. Verma, "Extreme gradient boost classifier based credit card fraud detection model," in *Proc. Int. Conf. Device Intell., Comput. Commun. Technol., (DICCT)*, Mar. 2023, pp. 500–504, doi: 10.1109/DICCT56244.2023.10110188.

[88] I. D. Mienye and Y. Sun, "A deep learning ensemble with data resampling for credit card fraud detection," *IEEE Access*, vol. 11, pp. 30628–30638, 2023, doi: 10.1109/access.2023.3262020.
[89] J. Novakovic and S. Markovic, "Performance of support vector machine in imbalanced data set," in *Proc. 19th Int. Symp. Infoteh-Jahorina (INFOTEH)*, Mar. 2020, pp. 1–5, doi: 10.1109/infoteh48170.2020.9066276.
[90] G. Zioviris, K. Kolomvatsos, and G. Stamoulis, "Credit card fraud detection using a deep learning multistage model," *J. Supercomput.*, vol. 78, no. 12, pp. 14571–14596, Aug. 2022, doi: 10.1007/s11227-022-04465-9.
[91] J. K. Afriyie, K. Tawiah, W. A. Pels, S. Addai-Henne, H. A. Dwamena, E. O. Owiredu, S. A. Ayeh, and J. Eshun, "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," *Decis. Anal. J.*, vol. 6, Mar. 2023, Art. no. 100163, doi: 10.1016/j.dajour.2023.100163.
[92] 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," *J. Inf. Secur. Appl.*, vol. 78, Nov. 2023, Art. no. 103618, doi: 10.1016/j.jisa.2023.103618.
[93] J. F. Roseline, G. Naidu, V. S. Pandi, S. A. A. Rajasree, and D. N. Mageswari, "Autonomous credit card fraud detection using machine learning approach," *Comput. Electr. Eng.*, vol. 102, Sep. 2022, Art. no. 108132, doi: 10.1016/j.compeleceng.2022.108132.
[94] S. Bakhtiari, Z. Nasiri, and J. Vahidi, "Credit card fraud detection using ensemble data mining methods," *Multimedia Tools Appl.*, vol. 82, no. 19, pp. 29057–29075, Aug. 2023, doi: 10.1007/s11042-023-14698-2.
[95] B. Xu, Y. Wang, X. Liao, and K. Wang, "Efficient fraud detection using deep boosting decision trees," *Decis. Support Syst.*, vol. 175, Dec. 2023, Art. no. 114037, doi: 10.1016/j.dss.2023.114037.
[96] A. D. Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in *Proc. IEEE Symp. Ser. Comput. Intell.*, Dec. 2015, pp. 159–166, doi: 10.1109/SSCI.2015.33.
[97] (Oct. 2, 2022). *IEEE Computational Intelligence Society*. IEEE-CIS Fraud Detection. Kaggle. [Online]. Available: https://www.kaggle.com/c/ieee-fraud-detection/data
[98] I.-C. Yeh and C.-H. Lien, "The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients," *Exp. Syst. Appl.*, vol. 36, no. 2, pp. 2473–2480, 2009.
[99] (2009). *UCSD: University of California, San Diego Data Mining Con-test 2009.* [Online]. Available: https://www.cs.purdue.edu/commugrate/data/credit_card/
[100] Z. A. Huang, Y. Sang, Y. Sun, and J. Lv, "A neural network learning algorithm for highly imbalanced data classification," *Inf. Sci.*, vol. 612, pp. 496–513, Oct. 2022, doi: 10.1016/j.ins.2022.08.074.
[101] I. D. Mienye and Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," *IEEE Access*, vol. 10, pp. 99129–99149, 2022, doi: 10.1109/access.2022.3207287.
[102] Z. Wang, P. Jia, X. Xu, B. Wang, Y. Zhu, and D. Li, "Sample and feature selecting based ensemble learning for imbalanced problems," *Appl. Soft Comput.*, vol. 113, Dec. 2021, Art. no. 107884, doi: 10.1016/j.asoc.2021.107884.

SURAYA NURAIN KALID was born in Kuala Terengganu, Terengganu, Malaysia, in 1990. She received the Diploma, bachelor's, and master's degrees in IT from Multimedia University, Malaysia, in 2010, 2014, and 2020, respectively. From 2014 to 2016, she was an Assistant Manager of core network, assisting in upgrading IP infrastructure for Telekom Malaysia's (TM) backbone. Then, she was assigned to TM's subsidiary Fiberail as an IT Executive with the Network Management Centre. Since November 2016, she has been a Lecturer with Multimedia University. Her research interests include artificial intelligence, data mining, and software development.

KOK-CHIN KHOR received the B.Sc. and M.Sc. degrees from Universiti Putra Malaysia, Seri Kembangan, Malaysia, and the Ph.D. degree from Multimedia University, Cyberjaya, Malaysia. Apart from teaching, he is also a certified Instructor of Cisco professional networking courses. Prior to joining the academic sector, he was a web programmer and a system analyst. His research interests include data mining and machine learning. He was the chair of several special sessions of international research conferences and a reviewer of several journals and conferences.

KENG-HOONG NG is an Assistant Professor with the Lee Kong Chian Faculty of Engineering Science, Universiti Tunku Abdul Rahman. He has been in the education sector as an academic staff, since 2002. He enjoys lecturing, providing training (CCNA professional course), and researching. He likes to travel, exercise (badminton), and find undervalued stocks in his spare time for his interest. His research interests include data clustering, data classification, bioinformatics, and computer networks.

GEE-KOK TONG is a Lecturer with the Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia. He has been conducting research work in financial econometric time series, such as the vector error correction model, GARCH modeling, extreme value theory, copula, value-at-risk, conditional value-at-risk, risk-adjusted performance measures, portfolio management, and Monte Carlo simulation.