computers

MDPI

Systematic Review

A Systematic Review of Machine Learning in Credit Card Fraud Detection Under Original Class Imbalance

Nazerke Baisholan [1,2,*], J. Eric Dietz [3], Sergiy Gnatyuk [4,5], Mussa Turdalyuly [2,6,7,*], Eric T. Matson [3] and Karlygash Baisholanova [1,*]

[1] Faculty of Information Technology, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan
[2] Software Engineering Department, International Engineering and Technological University, Almaty 050060, Kazakhstan
[3] Department of Computer and Information Technology, Purdue University, West Lafayette, IN 47907, USA; jedietz@purdue.edu (J.E.D.); ematson@purdue.edu (E.T.M.)
[4] Faculty of Computer Science and Technology, State University "Kyiv Aviation Institute", 03058 Kyiv, Ukraine; s.gnatyuk@kai.edu.ua
[5] State Scientific and Research Institute of Cybersecurity Technologies and Information Protection, 03142 Kyiv, Ukraine
[6] School of Digital Technologies, Narxoz University, Almaty 050035, Kazakhstan
[7] School of Engineering and Information Technologies, Eurasian Technological University, Almaty 050012, Kazakhstan
* Correspondence: nbaishol@purdue.edu (N.B.); m.turdalyuly@gmail.com (M.T.); baisholanova.k@gmail.com (K.B.)

Abstract

Credit card fraud remains a significant concern for financial institutions due to its low prevalence, evolving tactics, and the operational demand for timely, accurate detection. Machine learning (ML) has emerged as a core approach, capable of processing large-scale transactional data and adapting to new fraud patterns. However, much of the literature modifies the natural class distribution through resampling, potentially inflating reported performance and limiting real-world applicability. This systematic literature review examines only studies that preserve the original class imbalance during both training and evaluation. Following PRISMA 2020 guidelines, strict inclusion and exclusion criteria were applied to ensure methodological rigor and relevance. Four research questions guided the analysis, focusing on dataset usage, ML algorithm adoption, evaluation metric selection, and the integration of explainable artificial intelligence (XAI). The synthesis reveals dominant reliance on a small set of benchmark datasets, a preference for tree-based ensemble methods, limited use of AUC-PR despite its suitability for skewed data, and rare implementation of operational explainability, most notably through SHAP. The findings highlight the need for semantics-preserving benchmarks, cost-aware evaluation frameworks, and analyst-oriented interpretability tools, offering a research agenda to improve reproducibility and enable effective, transparent fraud detection under real-world imbalance conditions.

Keywords: credit card fraud detection; financial fraud; imbalanced classification; machine learning; systematic literature review; Kitchenham approach; fraud evaluation metrics; XAI; SHAP

check for updates

Academic Editor: Paolo Bellavista

Received: 14 August 2025
Revised: 12 September 2025
Accepted: 12 October 2025
Published: 15 October 2025

Citation: Baisholan, N.; Dietz, J.E.; Gnatyuk, S.; Turdalyuly, M.; Matson, E.T.; Baisholanova, K. A Systematic Review of Machine Learning in Credit Card Fraud Detection Under Original Class Imbalance. *Computers* **2025**, *14*, 437. https://doi.org/10.3390/computers14100437

1. Introduction

Financial fraud refers to the use of deceptive or illegal practices to obtain unauthorized financial gains [1]. This broad category includes credit card fraud, tax evasion, financial statement manipulation, money laundering, and synthetic identity creation [2]. Among these, credit card fraud remains one of the most pervasive and damaging forms, affecting the global financial landscape, costing businesses, consumers, and governments billions of dollars annually. The scope and sophistication of fraud have escalated sharply in recent years, driven by the widespread adoption of digital technologies and the emergence of advanced tools like generative artificial intelligence.

According to the European Central Bank and European Banking Authority, over 70% of fraudulent payment transactions across the EU in recent years were associated with digital channels, including phishing attacks, account takeovers, and card-not-present (CNP) fraud scenarios [3]. This trend underscores the growing exposure of electronic payment systems to cyber-enabled threats. In a global context, the Nilson Report highlights a significant imbalance in fraud distribution. Although U.S.-issued cards accounted for just 25.29% of global transaction volume in 2023, they were responsible for a disproportionate 42.32% of total card fraud losses worldwide [4]. This reflects a persistent vulnerability within U.S. payment infrastructure, where fraud losses remain concentrated despite widespread implementation of Europay, Mastercard, and Visa (EMV) technology and stronger authentication measures. Total global fraud losses exceeded $34 billion in 2023, marking the highest level in the past seven years and illustrating the continued escalation of sophisticated, cross-border financial crime. The Entrust 2025 Identity Fraud Report further emphasizes the exponential growth of AI-assisted fraud, with digital document forgeries increasing 244% year-over-year and deepfakes now comprising 40% of biometric fraud incidents [5].

As the digital economy expands, so too does the complexity of fraud, necessitating advanced, scalable, and explainable detection systems. Traditional manual detection mechanisms are no longer sufficient. In response, the financial industry is increasingly turning toward machine learning (ML) and explainable artificial intelligence (XAI) technologies to enhance fraud detection and resilience [6–8]. These intelligent systems are capable of processing high-dimensional data, identifying complex fraud patterns, and offering interpretable insights critical for compliance and trustworthiness.

Despite the growing body of research in financial fraud detection, many existing review studies have addressed specific domains such as online banking fraud [9,10], payment card misuse [11,12], or healthcare fraud [13,14]. At the same time, relatively few have provided a consolidated view focusing exclusively on credit card fraud detection through ML approaches [1,15,16]. Furthermore, a significant portion of the reviewed works rely on oversampled or synthetically balanced datasets, which do not fully reflect real-world transaction environments where fraud cases remain extremely rare [17].

This review addresses a critical gap by synthesizing machine learning approaches for credit card fraud detection under the original class distribution, without applying over- or undersampling during training or evaluation, and drawing on recent studies published between 2019 and 2025. Using a Kitchenham-style protocol [18], we systematically screened and extracted evidence from the literature, retaining only studies that reported at least Precision and Recall alongside a minimum set of evaluation metrics to enable comparability. The analysis quantifies patterns in dataset usage, model choice, and evaluation practices under true class priors, and assesses the extent to which interpretability is incorporated to support operational deployment.

The remainder of this article is structured as follows. Section 2 describes the review methodology, including the search strategy, screening protocol, inclusion and exclusion criteria, and quality assessment procedures. Section 3 presents the results of the systematic review, organized thematically to address dataset usage, algorithmic approaches, evaluation practices, and the integration of interpretability techniques. Section 4 offers a critical discussion of the implications for both research and practice. Section 5 outlines the study's limitations and threats to validity. Section 6 concludes with key takeaways and directions for future research.

2. Methodology

This research adopts a systematic literature review (SLR) framework to collect, evaluate, and synthesize existing research on ML methods for credit card fraud detection. The purpose of using an SLR approach is to ensure that the review process is transparent, unbiased, and replicable, allowing for a structured analysis of relevant literature that directly addresses the defined research questions. By aggregating high-quality, peer-reviewed sources, this review aims to offer evidence-based insights and minimize potential researcher bias throughout the selection and interpretation phases.

The methodology employed in this review follows the guidelines outlined by Kitchenham and colleagues [18], which organize the SLR process into three primary phases: planning the review, conducting the review, and reporting the review. Each phase involves specific tasks that ensure the integrity and comprehensiveness of the review process, as illustrated in Figure 1.

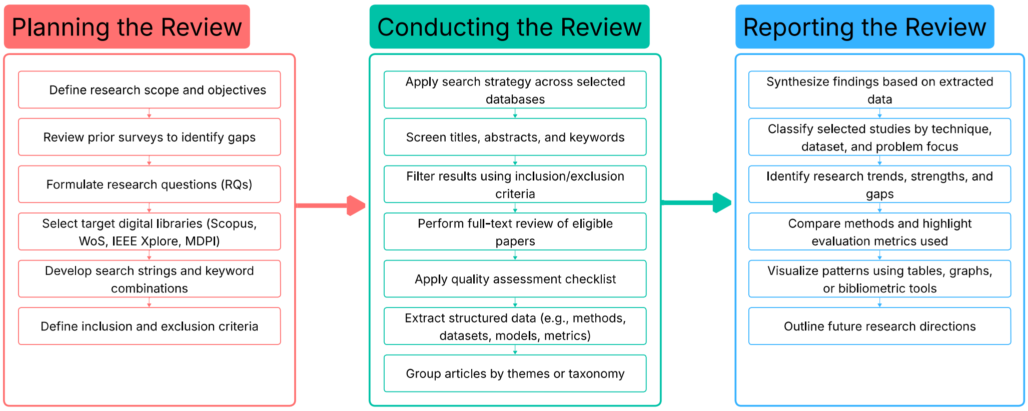

Planning the Review
Define research scope and objectives
Review prior surveys to identify gaps
Formulate research questions (RQs)
Select target digital libraries (Scopus, WoS, IEEE Xplore, MDPI)
Develop search strings and keyword combinations
Define inclusion and exclusion criteria
Conducting the Review
Apply search strategy across selected databases
Screen titles, abstracts, and keywords
Filter results using inclusion/exclusion criteria
Perform full-text review of eligible papers
Apply quality assessment checklist
Extract structured data (e.g., methods, datasets, models, metrics)
Group articles by themes or taxonomy
Reporting the Review
Synthesize findings based on extracted data
Classify selected studies by technique, dataset, and problem focus
Identify research trends, strengths, and gaps
Compare methods and highlight evaluation metrics used
Visualize patterns using tables, graphs, or bibliometric tools
Outline future research directions

Figure 1. Systematic literature review strategy based on Kitchenham's guidelines.

The primary objective of this review is to provide a comprehensive synthesis of ML-based approaches for credit card fraud detection, with particular emphasis on:

- Techniques that address highly imbalanced datasets;
- Models that prioritize Recall, area under the precision–recall curve (AUC-PR), and explainability;
- The use of Shapley additive explanations (SHAP) and other XAI methods for fraud interpretability.

This study does not focus on generic financial fraud domains such as insurance, telecom, or health fraud, but instead narrows the scope to transaction-level fraud in credit card systems.

As part of the review planning phase, we conducted an exploratory analysis of existing literature surveys in the domain of financial fraud detection. We identified several review articles related to credit card fraud detection [1,2,15,19]. However, none of them emphasized studies that retain the original class imbalance typically present in real-world fraud datasets. This review aims to address that gap by focusing on models evaluated under naturally imbalanced conditions, where fraud remains a rare event.

Moreover, prior surveys commonly rely on general metrics and area under the receiver operating characteristic curve (AUC-ROC), which may not adequately reflect performance in imbalanced settings. In contrast, this review prioritizes AUC-PR over AUC-ROC, which is more appropriate for evaluating fraud detection models. Additionally, the role of model interpretability, crucial in financial applications, has received limited attention in prior reviews [1,2,15,16].

This study addresses these gaps by focusing on ML-based credit card fraud detection methods that preserve dataset imbalance, apply suitable evaluation metrics, and incorporate explainability techniques such as SHAP.

To guide the structure and focus of this systematic review, we formulated a set of research questions (RQs). These questions were designed to capture the core aspects of ML-based credit card fraud detection, with a particular emphasis on real-world challenges such as class imbalance, metric selection, and model interpretability. Table 1 outlines the four main research questions along with their corresponding motivations, which serve as the foundation for article selection, categorization, and analysis in subsequent sections.

Table 1. Research questions and motivations.

No.	RQ	Motivation
1	Which publicly available or private datasets are most frequently used in credit card fraud detection studies that retain the original class imbalance without applying resampling techniques?	To characterize the empirical basis of non-resampled credit-card fraud research by identifying which public, private, and synthetic datasets are used, how representative they are of production class skews, and why authors choose to preserve imbalance.
2	What types of machine learning algorithms are most frequently applied in credit card fraud detection studies that retain the original class imbalance?	To identify the range of ML approaches, such as ensemble methods, deep learning models, and hybrid architectures, used to detect fraud in transaction data.
3	In studies that retain the original class imbalance, which evaluation metrics are prioritized, and how frequently are Recall and AUC-PR favored over Accuracy or AUC-ROC?	To assess whether reported metrics align with rare-event detection objectives by measuring the emphasis on Recall/Precision/F1-score and AUC-PR versus Accuracy and AUC-ROC.
4	How are interpretability techniques operationalized in non-resampled credit-card fraud studies to explain predictions and support operational deployment?	To determine how explainability is integrated when imbalance is retained, what tools are used, whether explanations are provided at global and case levels for governance and triage, and what barriers limit adoption.

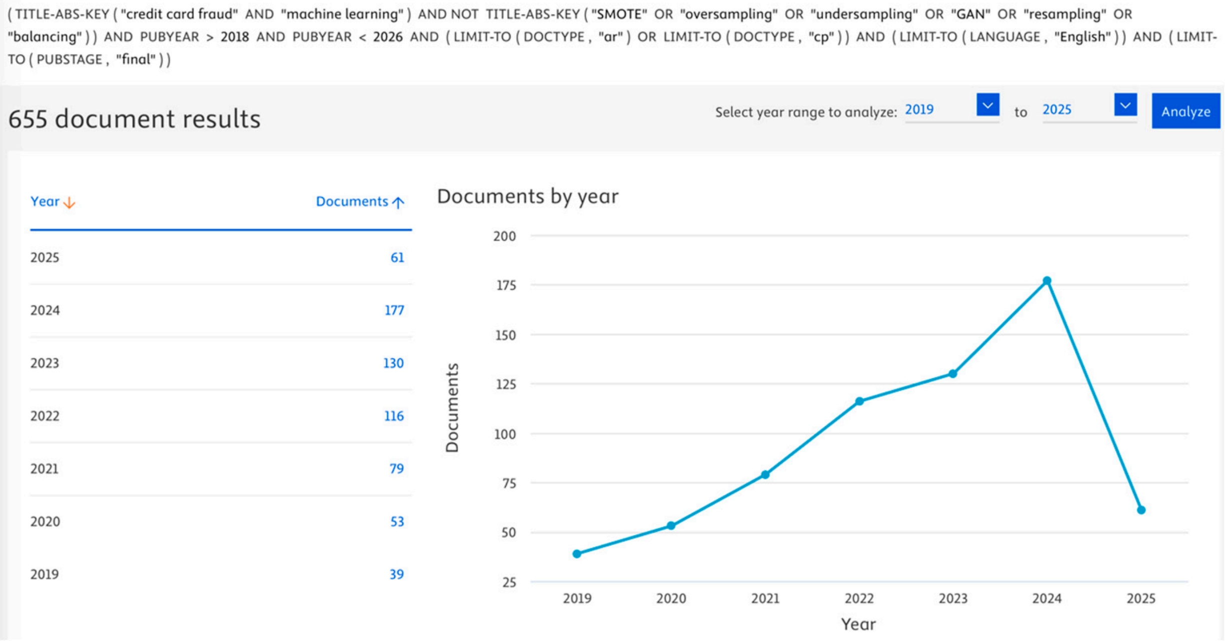
The search process was guided by a structured keyword query combining domain-specific and technical terms. Boolean logic (AND, OR) was used to form robust search strings across multiple digital libraries. The detailed criteria and sources are summarized in Table 2.

The systematic search was conducted in July 2025 across four major scholarly databases: Scopus, Web of Science, IEEE Xplore, and MDPI. The search strategy and eligibility criteria (outlined in Table 2) were developed in accordance with the PRISMA 2020 guidelines [20] to identify relevant studies on machine learning applications in credit card fraud detection. Study screening involved title and abstract filtering, exclusion of irrelevant publication types, and full-text review of eligible articles.

To illustrate the temporal distribution of retrieved publications, Figure 2 presents results from the Scopus database as a representative example. The number of publications related to credit card fraud detection using machine learning increased consistently from 2019 to 2024, indicating growing academic and industrial interest in combating fraud with AI-driven techniques. The peak in 2024 reflects heightened attention to this domain, while preliminary data for 2025 suggest sustained research activity.

Table 2. Inclusion criteria, search strategy, and data sources used in the review.

Aspect	**Details**
Time Frame	Publications between January 2019 and July 2025 to capture recent advancements and trends.
Language	Only studies published in English were considered.
Publication Types	Peer-reviewed journal articles and conference papers.
Relevance Criteria	Focus on credit card fraud detection, machine learning, or AI-based detection techniques.
Technological Scope	Inclusion of studies applying machine learning, deep learning, or ensemble methods to fraud detection problems.
Exclusion Criteria	Studies lacking empirical evaluation, vague methodology, or unrelated to credit card fraud were excluded. Additionally, studies with unsubstantiated or contradictory performance results that could not be validated through their provided data or figures were omitted. Furthermore, studies were excluded if they used unnamed or proprietary datasets without a clear description, hindering reproducibility and transparency.
Search Keywords	(“credit card fraud” OR “transaction fraud” OR “card-not-present fraud”) AND (“machine learning” OR “deep learning” OR “AI” OR “XAI”)
Digital Libraries Used	Scopus, Web of Science (WoS), IEEE Xplore, MDPI



```
( TITLE-ABS-KEY ( "credit card fraud"  AND  "machine learning" )  AND NOT  TITLE-ABS-KEY ( "SMOTE"  OR  "oversampling"  OR  "undersampling"  OR  "GAN"  OR  "resampling"  OR
"balancing" ) )  AND  PUBYEAR  >  2018  AND  PUBYEAR  <  2026  AND  ( LIMIT-TO ( DOCTYPE ,  "ar" )  OR  LIMIT-TO ( DOCTYPE ,  "cp" ) )  AND  ( LIMIT-TO ( LANGUAGE ,  "English" ) )  AND  ( LIMIT-
TO ( PUBSTAGE ,  "final" ) )
```

655 document results

Select year range to analyze: 2019 to 2025

Year ↓	Documents ↑
2025	61
2024	177
2023	130
2022	116
2021	79
2020	53
2019	39

Figure 2. Year-wise distribution of retrieved Scopus articles (2019–2025) for studies on “credit card fraud” and “machine learning,” excluding sampling-based methods.

The study selection process, along with the number of records included and excluded at each stage, is illustrated in the PRISMA 2020 flow diagram (Figure 3) and is further described in the Section 3.

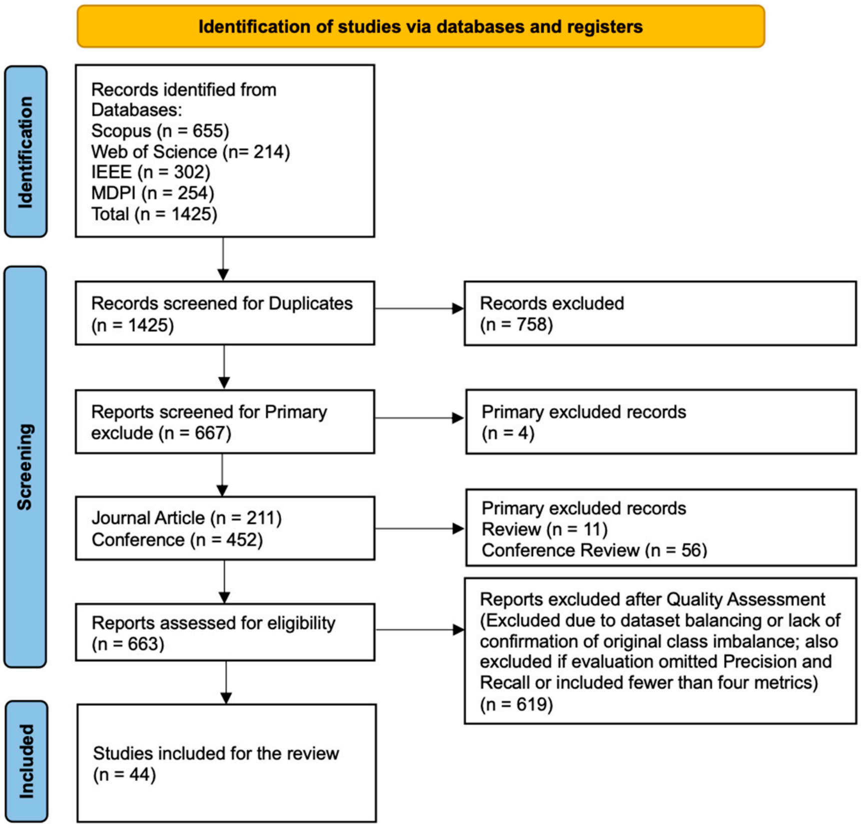

Identification of studies via databases and registers
Identification
Records identified from Databases: Scopus (n = 655) Web of Science (n= 214) IEEE (n = 302) MDPI (n = 254) Total (n = 1425)
Screening
Records screened for Duplicates (n = 1425)
Records excluded (n = 758)
Reports screened for Primary exclude (n = 667)
Primary excluded records (n = 4)
Journal Article (n = 211) Conference (n = 452)
Primary excluded records Review (n = 11) Conference Review (n = 56)
Reports assessed for eligibility (n = 663)
Reports excluded after Quality Assessment (Excluded due to dataset balancing or lack of confirmation of original class imbalance; also excluded if evaluation omitted Precision and Recall or included fewer than four metrics) (n = 619)
Included
Studies included for the review (n = 44)

Figure 3. PRISMA 2020 flow diagram illustrating the study selection process.

3. Results

3.1. Study Selection

The systematic search conducted across four major digital libraries: Scopus, Web of Science, IEEE Xplore, and MDPI yielded a total of 1425 records (Scopus: 655; Web of Science: 214; IEEE: 302; MDPI: 254). After removing 758 duplicates and irrelevant entries, 667 records remained for further screening.

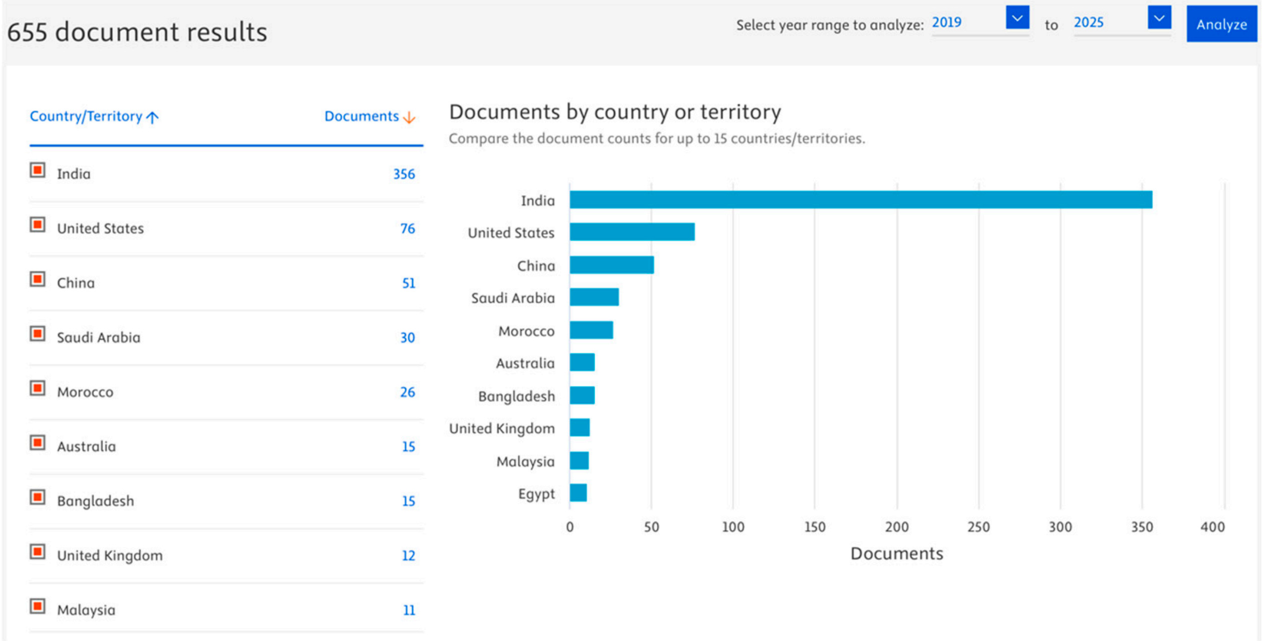
An additional four records were excluded at the preliminary stage due to incorrect metadata or inaccessible full texts. The remaining 663 records (211 journal articles and 452 conference papers) were subjected to full-text assessment against the inclusion and exclusion criteria (outlined in Table 2).

Following a rigorous quality assessment, 619 records were excluded due to:

- Resampling or synthetic augmentation (e.g., SMOTE/ADASYN, random over/ undersampling, GAN-based generation);
- Absence of confirmation that the original class imbalance was retained;
- Insufficient evaluation: studies that did not report both Recall and Precision, or reported fewer than four evaluation metrics.
- Lack of dataset transparency, such as unnamed or proprietary datasets without a clear description.

The final set of included studies comprised 44 peer-reviewed publications, including 26 conference papers and 18 journal articles. These studies specifically addressed credit card fraud detection using machine learning while retaining the original class imbalance and reporting at least four evaluation metrics, including both Recall and Precision. A visual summary of the selection process is shown in Figure 3, following the PRISMA 2020 flow diagram.

To complement the PRISMA analysis and provide an overview of the research landscape, Figures 4 and 5 summarize the country-wise and document-type distributions of the retrieved Scopus records (n = 655), prior to quality assessment and inclusion.



655 document results

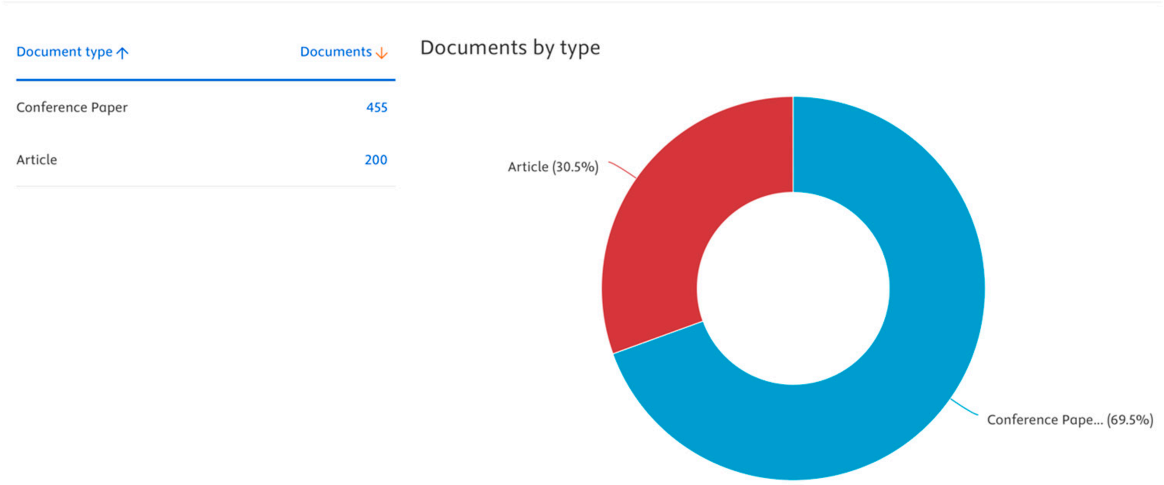
Select year range to analyze: 2019 to 2025 Analyze

Country/Territory ↑	Documents ↓
☑ India	356
☑ United States	76
☑ China	51
☑ Saudi Arabia	30
☑ Morocco	26
☑ Australia	15
☑ Bangladesh	15
☑ United Kingdom	12
☑ Malaysia	11

Documents by country or territory

Compare the document counts for up to 15 countries/territories.

Figure 4. Country-wise distribution of retrieved Scopus documents (2019–2025) related to credit card fraud and machine learning.



Document type ↑	Documents ↓
Conference Paper	455
Article	200

Documents by type

Figure 5. Distribution of Scopus documents by publication type: conference papers and journal articles.

As shown in Figure 4, India led in the number of publications (n = 356), followed by the United States (n = 76) and China (n = 51), indicating regional research concentration in fraud detection using ML.

In terms of publication type, Figure 5 illustrates that conference proceedings constituted the majority (69.5%, n = 455), while journal articles accounted for 30.5% (n = 200). This dominance of conference papers suggests rapid dissemination of evolving methodologies in this fast-paced research domain.

3.2. Results Organized by Research Questions

This section presents the findings of the systematic literature review, systematically organized by the four predefined Research Questions (RQs). Each subsection synthesizes the relevant evidence extracted from the included studies.

3.2.1. RQ1: Which Publicly Available or Private Datasets Are Most Frequently Used in Credit Card Fraud Detection Studies That Retain the Original Class Imbalance Without Applying Resampling Techniques

A review of the 44 included studies shows that a relatively small set of benchmark datasets dominates research in credit card fraud detection when the original class distribution is preserved. These datasets span real-world and synthetic sources and vary in size, feature semantics, and fraud prevalence, but all are used without oversampling or undersampling during training or evaluation.

Publicly available datasets remain the primary empirical basis. Particularly, the European Credit Card Fraud dataset is used in 32 of 44 studies (72.7%) and continues to serve as the benchmark due to its accessibility, realistic prevalence (~0.17% fraud), and long-standing role in comparative evaluation. The following most frequently used public sources are the IEEE-CIS/Vesta e-commerce dataset, employed by three of the 44 studies (6.8%), and the Credit Card Transactions Fraud Detection Dataset (Sparkov Dataset), also used by three of the 44 studies (6.8%). IEEE-CIS provides large-scale, multi-table online payment records with identity and device attributes. In contrast, Sparkov Dataset offers a synthetic, privacy-preserving transaction stream with documented class prevalence suitable for rare-event benchmarking.

Private datasets appear less often because of access constraints but add functional heterogeneity: a Brazilian bank dataset features in three studies (6.8%), the UCSD–FICO 2009 dataset and the Greek financial institution dataset each appear in only one study (2.3%). Among synthetic sources beyond Sparkov, BankSim and PaySim each appear in two studies (4.5%), and IBM TabFormer is used in one study (2.3%). These synthetic resources support controlled experimentation and privacy protection, but inevitably abstract away operational semantics and merchant- or customer-level context. Table 3 summarizes the key datasets identified in the reviewed studies, including their source, size, fraud ratio, availability, and frequency of use in the literature.

Following Table 3, authors commonly justify retaining the native class skew to mirror operational conditions and to ensure that reported performance reflects deployment realities. Many note that training on the original distribution avoids synthetic noise and preserves decision boundaries better than resampling, while others caution that oversampling, particularly before train-validation splitting, can induce information leakage and inflate reported results.

While concentration on the European dataset has facilitated reproducibility and comparability, heavy reliance on a single anonymized benchmark risks narrowing insights into robustness and transferability. Future progress would benefit from public, semantics-preserving benchmarks that balance privacy with interpretability, thereby enabling stronger XAI, more realistic threshold selection, and more explicit cross-method comparisons.

Table 3. Summary of Commonly Used Fraud-Detection Datasets and Their Characteristics.

Dataset Name	Transactions	Fraud %	Availability	Notes	Studies
European Credit Card Fraud Dataset [21,22]	284,807	0.172	Public	PCA features V1–V28; widely used benchmark	[7,23–52]
IEEE-CIS Fraud Detection Dataset [53]	1 M+	3.5	Public	PCA features V1–V339; transaction + identity attributes	[49,54,55]
Credit Card Transactions Fraud Detection Dataset (Sparkov Dataset) [56]	1,048,575	2.83	Public (synthetic)	1000 customers, 800 merchants	[57–59]
Brazilian Bank's Dataset	374,823	3.74	Private	Real bank data, 17 features	[28,60,61]
BankSim Dataset [62,63]	594,643	1.21	Public (synthetic)	Agent-based simulator; realistic fraud injection	[64,65]
PaySim Dataset [66,67]	6,362,620	0.13	Public (synthetic)	Mobile money simulation; transaction-level features	[68,69]
UCSD-FICO Data Mining Contest 2009 Dataset	94,683	2.21	Private	20 anonymized features; 73,729 cards	[24]
IBM Tabformer Dataset [70]	24 M	0.124	Public (synthetic)	6139 merchants, 100,343 merchants	[59]
Credit Card Fraud Data [71]	14,446	11.97	Public (synthetic)	11 features spanning domains	[72]
Greek Bank's Dataset	6,915,699	<0.1	Private	Real bank data; merchant/person subsets	[73]

3.2.2. RQ2: What Types of Machine Learning Algorithms Are Most Frequently Applied in Credit Card Fraud Detection Studies That Retain the Original Class Imbalance

Across the 44 included studies that preserve the original class distribution, a wide range of ML algorithms have been employed, spanning classical supervised learners, tree-based ensembles, deep and hybrid models, and unsupervised anomaly detection methods. In these works, class imbalance is addressed primarily through model-level strategies, such as class/cost weighting [7,23,26,61,69], custom loss functions [29], and decision-threshold tuning [7,26,48,49], rather than oversampling or undersampling. Tree-based ensembles are the most prevalent family, appearing in thirty-two of the forty-four studies (72.7%). Linear/margin models (primarily Logistic Regression and Support Vector Machines (SVM)) are also widespread, followed by single decision trees and neural/deep architectures. Secondary families include ensemble meta-strategies such as bagging/stacking and probabilistic methods (Naïve Bayes/Gaussian Mixture Models (GMM)). Instance-based methods (k-NN) and unsupervised/anomaly detectors (Isolation Forest, Local Outlier Factor (LOF), One-Class SVM, Autoencoders) are less common. Figure 6 visualizes the distribution of algorithm families, while the detailed per-algorithm counts and references are presented across Tables 4–7, grouped by tree-based ensembles, linear/probabilistic models, neural network approaches, and emerging methods.

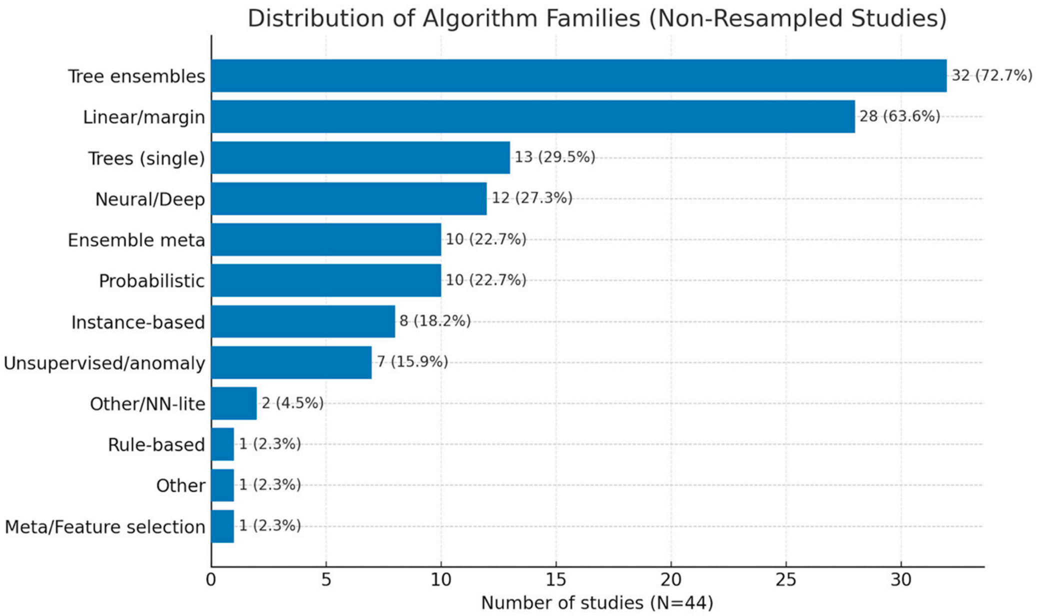


Figure 6. Distribution of algorithm families in credit card fraud detection studies that retain the original class imbalance.

Table 4. Tree-based and ensemble methods in credit card fraud detection without resampling.

Technique	**Description**	**Studies**
Random Forest	Ensemble of decision trees with bootstrap sampling; robust on tabular data; supports class weighting.	24 studies [7,24–27,30–33,35–37,39,43,46–48,50,51,54,57,59,64,72]
Decision Tree	Greedy recursive partitioning using impurity criteria yields interpretable rules and serves as a baseline or base learner in ensembles.	13 studies [7,24,30,33,34,36,37,49,51,54,57,64,72]
XGBoost	Regularized gradient-boosted trees with shrinkage and column sampling; strong performance on structured data.	12 studies [7,29,31,39,40,42,50,51,54,59,69,72]
LightGBM	Histogram-based gradient boosting with leaf-wise growth strategy; efficient and scalable for large tabular datasets.	10 studies [7,24,35,42,49,51,52,55,59,69]
Bagging/ Stacking/Voting	Meta-ensembles combining multiple base learners via averaging or a meta-model to improve stability and Accuracy.	10 studies [7,27,28,33,40,49,50,55,61,68]
Gradient Boosting	Additive tree boosting that iteratively fits residuals; umbrella term for Gradient Boosting Machine (GBM)-style learners.	6 studies [7,23,27,51,54,68]
AdaBoost	Boosting weak learners via reweighting; emphasizes previously misclassified samples to reduce bias.	6 studies [23,25,32,33,50,54]
CatBoost	Ordered boosting with target statistics for categorical features; reduces target leakage and overfitting.	5 studies [35,47,51,59,69]
Other Tree/Rule Ensembles (LiteMORT, RIPPER, RWN)	Specialized or niche ensemble methods applied in single studies.	3 studies [48,55,68]

Table 5. Linear and probabilistic models in credit card fraud detection without resampling.

Technique	Description	Studies
Logistic Regression	Linear probabilistic classifier with logistic link; widely used as a baseline and interpretable model.	19 studies [7,23,24,26,27,30–33,36,39,42,47,49,51,52,57–59]
SVM	Margin-based classifier (often with radial basis function (RBF) kernel); supports class weighting; effective in high-dimensional spaces.	14 studies [23,24,33,34,37,39,41,45,48,49,54,57,64,65]
Naïve Bayes	Generative classifier assuming conditional independence; simple but effective baseline for sparse data.	9 studies [7,24,26,33,34,48,49,51,57]
GMM	Probabilistic mixture of Gaussians used for clustering or density-based classification.	1 study [65]
One-Class SVM	Boundary method estimating the support of the majority class; applied for anomaly detection.	1 study [73]

Table 6. Neural network-based models in credit card fraud detection without resampling.

Feed-Forward Neural Network (MLP, ANN/FFNN)	Neural network with one or more hidden layers; capacity controlled by depth/width and regularization.	6 studies [7,26,28,35,52,64]
CNN	Convolutional neural networks applied to reshaped/tabular representations or learned embeddings.	5 studies [38,40,54,58,59]
LSTM	Recurrent neural network with gating for long-term dependencies; models sequential transaction patterns.	4 studies [28,38,40,59]
GRU	Neural reconstruction model for representation learning or anomaly detection via reconstruction error.	1 study [28]
Autoencoder	Neural reconstruction model for representation learning or anomaly detection via reconstruction error.	3 studies [35,58,73]
SL-SSNet	Representation learning with a downstream fraud classifier; leverages self-supervised pre-training.	1 study [51]

Across individual algorithms, Random Forest (RF) is most common (24 studies; 54.5%), followed by Logistic Regression (LR) (19 studies; 43.2%) and Support Vector Machines (SVM) (14 studies; 31.8%). Among the boosting variants, XGBoost appears in twelve studies (27.3%), LightGBM in ten (22.7%), and CatBoost in five (11.4%). This pattern reflects the field's preference for tree ensembles on tabular transactions, while retaining LR and SVM as strong, interpretable baselines.

Deep and hybrid models are reported in twelve studies (27.3%), including feed-forward neural networks (MLP/ANN), convolutional networks (CNNs), recurrent networks (LSTM/GRU), CNN–BiLSTM hybrids, graph neural networks (GNNs), and tabular attention models such as TabTransformer. These architectures are often paired with boosted trees in stacked ensembles and trained with weighted losses and regularization to mitigate bias toward the majority class. Unsupervised and semi-supervised anomaly detection remains less common (approximately 16% of studies), leveraging Isolation Forest, LOF, One-Class SVM, and Autoencoders to model legitimate behavior and flag deviations. Several papers integrate these detectors into hybrid pipelines in which a supervised model subsequently classifies anomalies identified in an unsupervised stage.

Table 7. Other and emerging approaches in credit card fraud detection without resampling.

Technique	Description	Studies
k-NN	Instance-based classifier; predicts class by majority vote of nearest neighbors.	8 studies [24,26,34,45,51,54,64,65]
ELM	Single-hidden-layer feed-forward network with random weights; fast training for classification tasks.	1 study [65]
Isolation Forest	Tree-based anomaly detection method; isolates rare points via recursive partitioning.	5 studies [41,44,45,64,73]
LOF	Density-based anomaly detection compares local reachability density to that of its neighbors.	3 studies [41,44,64]
Transformer/TabTransformer/TabNet	Attention-based/tabular-specific architectures that learn feature interactions and embeddings.	2 studies [40,49]
GNN	Graph neural networks are typically trained with class weighting or focal objectives to address imbalance, sometimes combined with tabular baselines in ensembles	1 study [59]
Bio-inspired/immune-based classifiers (AIS, ARO)	Immune-system and asexual-reproduction-inspired learners are used directly for fraud detection or rule optimization.	1 study [60]
RHSOFS	Rock hyrax swarm optimization-based wrapper for feature selection.	1 study [34]

Overall, the evidence indicates a clear methodological concentration on tree-based ensembles, particularly Random Forest and gradient-boosting methods, supported by interpretable linear baselines and a growing, though still modest, adoption of deep, attention-based, and graph-oriented architectures. This pattern reflects a pragmatic trade-off between achieving high recall-oriented performance under heavy class imbalance and retaining computational efficiency and interpretability for operational deployment.

3.2.3. RQ3: In Studies That Retain the Original Class Imbalance, Which Evaluation Metrics Are Prioritized, and How Frequently Are Recall and AUC-PR Favored over Accuracy or AUC-ROC

Selecting appropriate performance metrics is critical for financial fraud detection, where the positive class is exceedingly rare and the misclassification costs are highly asymmetric [2,74]. Under such conditions, overall Accuracy can be misleading. In contrast, metrics conditioned on the positive class, such as Recall, Precision, F1-score, and the area under the Precision–Recall curve (AUC-PR), provide a more faithful characterization of detection capability and support principled threshold selection for deployment [7,75–77]. Nevertheless, there is no universally accepted evaluation standard [2,16]. Recent studies report a heterogeneous mix of scalar and curve-based indicators, and only a small subset incorporates cost- or profit-based measures aligned with operational risk. Table 8 details each metric's formula and representative references.

Table 8. Performance evaluation metrics and formulas used in studies.

Metrics	Formula	Studies
Accuracy	$Accuracy = \frac{(TN+TP)}{(TN+FN+FP+TP)}$	[7,24–27,29–39,41–52,55,57,58,60,61,64,65,68,69,72,73]
Precision	$Precision = \frac{TP}{(TP+FP)}$	[7,23–52,54,55,57–61,64,65,68,69]
Recall (Sensitivity/TPR)	$Recall = \frac{TP}{(TP+FN)}$	[7,23–52,54,55,57–61,64,65,68,69]
F1-score	$F1 = 2 \times \frac{Recall \times Precision}{(Recall+Precision)}$	[7,23–28,30–52,54,55,57–59,64,65,68,69,72,73]
Specificity (TNR)	$Specificity = \frac{TN}{(TN+FP)}$	[34,40,50,60]
AUC-ROC	$AUC - ROC = \int_0^1 TPR\left(FPR^{-1}(x)\right) dx$	[23–26,28,36–38,40,42,48,52,54,55,61,64,65,68,69,72,73]
AUC-PR	$AUC - PR = \int_0^1 Precision(Recall)\ d(Recall)$	[7,28,38,59,64]
Matthews Correlation Coefficient (MCC)	$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$	[29,30,32–34,48,64,65,69]
Other metrics		[7,23,25–27,30,37,38,43,45,46,48,50–52,61,72,73]

In fraud detection, Accuracy quantifies the proportion of correctly classified transactions overall, while Precision reflects the reliability of positive (fraud) alerts. Recall (Sensitivity/TPR) measures the share of truly fraudulent transactions that are correctly identified, and Specificity (TNR) measures the share of legitimate transactions correctly rejected [2]. Because Precision and Recall move in opposite directions as the decision threshold is varied, practitioners face the well-known precision–recall trade-off. A common scalar compromise is the F1-score, the harmonic mean of Precision and Recall [78]. In practice, varying the threshold and examining the induced sequence of confusion matrix constitutes parametric evaluation, which supports the selection of an operating point consistent with business constraints.

To ensure comparability, our inclusion criteria required studies to report both Precision and Recall, which are therefore present in all 44 studies (100%). Beyond these two, reporting practices concentrate on F1-score in 41 studies (93.2%) and Accuracy in 39 studies (88.6%). Curve-based metrics are less uniform. AUC-ROC appears in 21 studies (47.7%), whereas AUC-PR (or Average Precision) appears in only five studies (11.4%), typically where authors explicitly foreground severe class imbalance or adopt anomaly-detection framings [7,28,38,59,64]. Additional reported measures include Matthews Correlation Coefficient (MCC) (20.5%) and Specificity (TNR) (9.1%). At the same time, several papers provide confusion-matrix counts and, more occasionally, cost/profit or alert-rate summaries to support operational decision-making. Figure 7 provides an overview of how often each evaluation metric is reported across the included studies.

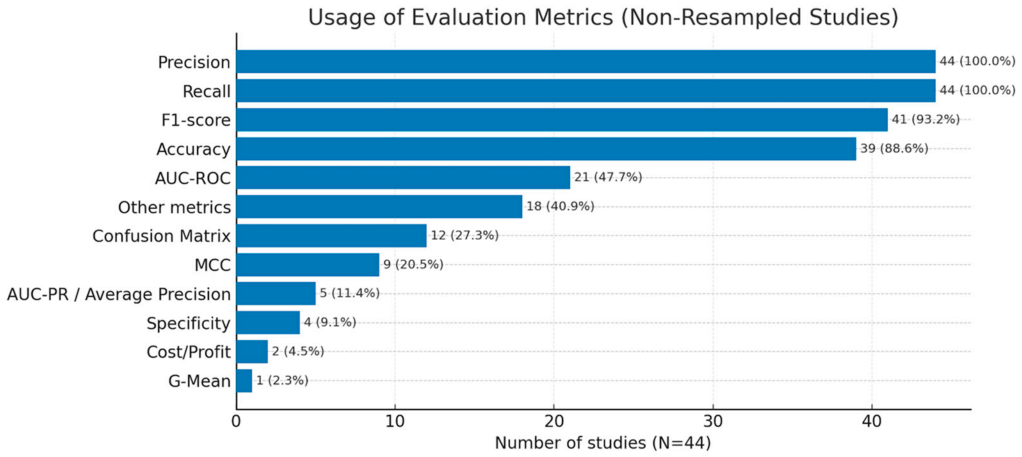


Figure 7. Reporting frequency of evaluation metrics in credit-card fraud detection studies that retain the original class imbalance.

Because Recall and Precision were mandatory for inclusion, all studies report them. However, the selection narratives are generally recall-oriented, balancing Recall with Precision or F1-score on the premise that missed fraud is costlier than extra alerts [25,31–33,36–39,42,50]. Accuracy is typically kept for completeness rather than for decision-making. Although many authors note that AUC-PR is more appropriate under severe imbalance, AUC-ROC persists as the default curve metric, mainly for historical and tooling reasons. While AUC-PR is treated as a co-primary indicator alongside Recall, F1-score [7,28,38,59,64]. Several studies explicitly caution that AUC-ROC can overstate performance in rare-event regimes [31,36–38,42,48,55,69,72]. Only a small minority connect statistical metrics to business outcomes, with explicit cost or profit reporting limited to total processing cost (TPC) in a cost-sensitive boosting setting [23] and total cost under cost-aware ensembles [62]. Taken together, the literature favors recall-conscious, precision–recall (PR)-aware evaluation, but there remains significant scope for standardization of reporting on AUC-PR, threshold-dependent summaries, confusion matrix disclosure, MCC, and cost-linked measures to support deployment decisions better.

3.2.4. RQ4: How Are Interpretability Techniques Operationalized in Non-Resampled Credit-Card Fraud Studies to Explain Predictions and Support Operational Deployment

Explainable artificial intelligence (XAI) aims to make the behavior of ML models transparent and understandable, particularly for end-users, regulatory authorities, organizational personnel, and other stakeholders [79]. In credit card fraud detection, this involves producing explanations that data scientists can validate, business users and investigators can act upon, and risk managers or regulators can audit. Interest in XAI has grown alongside the increasing deployment of ML in finance, driven by the need for trust, accountability, robustness, and compliance in high-stakes environments [80].

Within the reviewed literature, only two studies implemented operational interpretability, and both adopted SHAP [7,73]. In the FraudX AI framework [7], SHAP was applied to provide global and local attributions over PCA-anonymized features. Summary plots and feature-importance rankings supported auditor understanding and regulatory transparency, although the authors note that anonymization restricted semantic interpretation despite stable ranking of influential components.

In [73], an unsupervised, production-proximate study using a Greek bank's transaction data also integrated SHAP to generate both global and per-transaction explanations. Force and decision plots highlighted the drivers behind anomaly scores, aiding analysts in prioritization and reducing cognitive load. Interviews with monitoring staff indicated improved case review efficiency when SHAP outputs accompanied model alerts.

Across both implementations, SHAP is positioned as a model-agnostic interpretability layer serving two primary purposes:

1. Global governance—identifying the most influential features driving fraud predictions;
2. Local justification—providing case-level explanations to support analyst workflows and post hoc threshold adjustments.

Two additional studies did not implement XAI but explicitly identified it as planned work. Renganathan et al. [46] suggested that adding explainability could enhance model transparency, foster user trust, and mitigate bias. Jain et al. [58] likewise propose adopting SHAP and other model-agnostic methods to illuminate decision processes. Although not counted as operational XAI, these proposals reflect growing awareness of the need for explanation in governance and auditability, especially under original imbalanced distributions.

Overall, SHAP emerges as the predominant technique for operationalizing interpretability in non-resampled credit card fraud detection studies, valued for its model-agnostic design and ability to deliver both governance-level summaries and transaction-level justifications. However, adoption remains rare, constrained by: (i) semantic loss from anonymized feature spaces; (ii) lack of analyst-friendly explanation interfaces; and (iii) concerns about explanation stability under data drift and evolving fraud tactics. While SHAP dominates current implementations, other model-agnostic explainability techniques are emerging that could also be relevant for fraud detection. Local interpretable model-agnostic explanations (LIME), for instance, approximates complex model decisions with local surrogate models and have been applied in adjacent financial risk settings [81]. More recent approaches such as counterfactual explanations, which identify minimal feature changes that would alter a model's decision, are gaining attention as actionable tools for providing recourse in fraud detection. Diverse counterfactual explanations (DiCE) have been benchmarked alongside SHAP, LIME, and Anchors in ANN-based fraud models, showing complementary strengths and challenges [82]. Advances such as conformal prediction interval counterfactuals (CPICFs) further illustrate how individualized counterfactuals can reduce uncertainty and improve interpretability in transaction fraud datasets [83]. In parallel, attention-based interpretability mechanisms have been applied in online banking fraud detection, where hierarchical attention models can highlight both the most suspicious transactions and the most informative features, offering human-understandable explanations [84]. Recent studies further demonstrate the value of combining SHAP and LIME with human-in-the-loop oversight to support auditability and regulatory transparency in financial fraud detection frameworks [85]. Although not yet adopted in the non-resampled credit card fraud detection studies we reviewed, these techniques represent promising directions for extending explainability beyond SHAP in future research and deployment. This points to a concrete research agenda: (a) retain or document semantics-preserving feature mappings so attributions remain actionable; (b) provide threshold-dependent summaries alongside explanations to inform policy-setting; (c) assess the fidelity and robustness of explanations under temporal drift; and (d) evaluate the operational impact of XAI, measuring its influence on triage speed, alert quality, and analyst consensus in addition to standard statistical performance metrics.

4. Discussion

This review highlights the close interplay between dataset availability, model choice, evaluation practice, and explainability in credit card fraud detection studies that preserve the original class imbalance. The European Credit Card Fraud dataset overwhelmingly dominates as the benchmark of choice, appearing in two-thirds of the included studies, providing a shared basis for cross-paper comparability but also imposing constraints on interpretability due to PCA-based anonymization of features. Public datasets like IEEE-CIS and PaySim appear far less frequently. In contrast, private datasets from financial institutions remain valuable but are typically accessible only through collaborations or competition-based platforms [24,73,86]. Synthetic datasets such as BankSim and PaySim are leveraged where privacy concerns or controlled experimentation are priorities, though they often cannot fully capture evolving fraud behaviors.

Across the included studies, tree-based ensembles, notably RF and gradient boosting, are the most frequently applied algorithms. Their dominance is unsurprising for high-dimensional, heterogeneous tabular data, where they offer strong handling of mixed feature types, built-in support for class weighting, and competitive precision–recall trade-offs. Linear and margin-based baselines (LR, SVM) remain common for interpretability and benchmarking, while deep and hybrid architectures (CNN–BiLSTM, TabTransformer) are emerging in contexts where temporal or relational patterns are essential. Notably, most studies handle imbalance at the model level, using class weights, focal or modified loss functions, or threshold calibration, rather than applying resampling techniques such as SMOTE or undersampling. This reflects production realities, where preserving the actual fraud prior facilitates integration with downstream controls, prevents oversampling artifacts, and avoids leakage risks.

Evaluation practices in the reviewed literature are universally positive, class-oriented, with all studies reporting Recall and Precision. However, AUC-ROC remains more frequently reported than AUC-PR, despite the latter's greater relevance under extreme skew. For future research, AUC-PR should be treated as the primary evaluation metric, with AUC-ROC retained only as a supplementary indicator. Threshold-dependent reporting, such as confusion matrix counts, and operating-point precision/recall is also essential for transparency. Beyond statistical measures, cost-sensitive and profit-based metrics should be incorporated into future studies to reflect the asymmetric risks of false negatives and false positives in real-world deployment. Such adoption would significantly improve the comparability, operational realism, and policy relevance of fraud detection research.

Explainability remains underdeveloped: only two studies operationalize model interpretability, and both adopt SHAP. In these implementations, SHAP serves as a model-agnostic layer for generating global feature rankings to support governance and local transaction-level explanations to aid case review. While both report positive analyst feedback and improved triage efficiency, the utility of explanations is limited by anonymized feature spaces and the lack of user-friendly analyst interfaces. A small number of additional studies acknowledge the importance of explainability and propose integrating it in future work, underscoring its perceived but unrealized value in the field.

Taken together, these findings suggest several methodological and practical priorities for advancing credit card fraud detection research under non-resampled conditions:

1. Dataset diversification and transparency. Broaden the set of publicly available benchmark datasets, ideally with non-anonymized, privacy-preserving features to enable more actionable explainability and cross-study comparability;
2. Balanced model exploration. Maintain strong tree-based ensemble baselines while expanding evaluation of sequence, graph, and attention-based models in scenarios where temporal and relational structures are prominent;
3. Evaluation alignment with deployment needs. Prioritize PR-space metrics, threshold-specific performance, and cost/profit analysis to bridge model evaluation with operational decision-making;

4. Operational explainability. Preserve or securely map feature semantics to enable actionable attributions, design analyst-friendly explanation interfaces, and assess explanation robustness under data drift.

By integrating methodological rigor with operational realism, future studies can better align academic advances with the needs of fraud detection teams, regulators, and financial institutions, ultimately enhancing both the performance and trustworthiness of deployed systems.

5. Limitations and Threats to Validity

Several limitations should be acknowledged when interpreting this review. First, the evidence base is heavily skewed toward the European Credit Card Fraud dataset, which facilitates reproducibility but restricts external validity and constrains explainability due to PCA-based anonymization.

Second, our inclusion criteria, requiring explicit reporting of both Precision and Recall and at least four evaluation metrics, may bias the corpus toward studies with more comprehensive reporting practices, potentially excluding otherwise relevant work. Of the 663 reports assessed in full text, only 44 (6.6%) satisfied all quality requirements, reflecting the limited number of studies meeting these standards in the literature.

Third, heterogeneity in experimental protocols (data splits, cross-validation folds, temporal ordering, hyperparameter tuning strategies) limits the strict comparability of reported results across studies.

Fourth, the body of published studies may be subject to publication bias, as works with negative or inconclusive findings are less likely to be reported, which may limit the visibility of alternative approaches or experimental failures.

Finally, limited transparency in code, seeds, and preprocessing steps in many studies restricts the ability to replicate results, assess robustness, and conduct fair head-to-head comparisons.

These limitations also affect the generalizability of our findings. Heavy reliance on the European Credit Card dataset may restrict the applicability of observed methodological trends to other domains such as mobile payments or online banking, where transaction structures and fraud patterns differ. Likewise, inclusion and reporting biases may overrepresent well-documented or positive outcomes, overstating the maturity of the field. While the trends identified here are robust within the reviewed corpus, caution is warranted when generalizing these findings to the broader landscape of fraud detection research under class imbalance.

6. Conclusions

Financial fraud remains a persistent threat across banking and e-commerce, with losses amplified by the rarity and evolving nature of fraudulent transactions. Modern ML methods offer scalable screening of high-volume streams. Still, their operational value depends on faithfully reflecting real-world class imbalance, applying rigorous evaluation protocols, and delivering outputs that stakeholders can interpret and act upon.

This review systematically examined forty-four primary studies that retained the native class distribution. The evidence is grouped into four categories. Datasets: a small set of public benchmarks, most notably the European Credit Card Fraud dataset, anchors much of the literature, enabling replication but constraining semantic interpretability due to anonymization; private and synthetic datasets appear less frequently and are inconsistently documented. Algorithms: tree-based ensembles dominate non-resampled settings, with RL and SVM serving as standard baselines; deep and hybrid architectures (sequence, convolutional, attention-based, graph) are present but less common. Imbalance handling occurs primarily at the model/threshold level (class/cost weighting, modified losses, operating-point calibration) rather than via resampling. Evaluation: Precision and Recall are consistently reported, F1-score and Accuracy are frequent, while AUC-PR, more suitable for rare-event regimes, remains underused compared to AUC-ROC. Links between statistical performance and business outcomes are rare. Explainability: operational XAI adoption is limited; where present, SHAP is used for global governance and local triage, but anonymized features limit business-semantic insight.

The findings suggest clear priorities for future work. Data assets would benefit from semantics-preserving benchmarks or secure feature dictionaries to support interpretable modeling and policy translation. Methodologically, evaluation protocols must move beyond over-reliance on AUC-ROC. AUC-PR should be treated as the primary evaluation metric, complemented by threshold-specific reporting. In addition, cost- or profit-sensitive measures should become mandatory standards in fraud detection research under non-resampled conditions. This shift would better align academic benchmarks with the realities of operational deployment, where precision–recall trade-offs and asymmetric error costs dictate effectiveness.

Author Contributions: Conceptualization, N.B., S.G. and J.E.D.; methodology, N.B.; software, N.B.; validation, N.B., S.G., M.T. and J.E.D.; formal analysis, N.B. and M.T.; investigation, N.B.; resources, E.T.M. and K.B.; data curation, N.B.; writing—original draft preparation, N.B.; writing—review and editing, N.B., J.E.D. and E.T.M.; visualization, N.B.; supervision, S.G., J.E.D. and M.T.; project administration, M.T. and S.G.; funding acquisition, K.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP25794007).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AdaBoost	Adaptive Boosting
ADASYN	Adaptive Synthetic Sampling
AI	Artificial Intelligence
AIS	Artificial Immune System
ANN	Artificial Neural Network
ARO	Asexual Reproduction Optimization
AUC-PR	Area Under the Precision-Recall Curve
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
BiLSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
CNP	Linear dichroism
CPICFs	Conformal Prediction Interval Counterfactuals
DiCE	Diverse Counterfactual Explanations
DT	Decision Tree
ELM	Extreme Learning Machine
EMV	Europay, MasterCard, and Visa
EU	European Union
FFNN	Feed-Forward Neural Network
FN	False Negative
FP	False Positive
GAN	Generative Adversarial Network

GBM	Gradient Boosting Machine
GMM	Gaussian Mixture Model
GNN	Graph Neural Network
GRU	Gated Recurrent Unit
IBM	International Business Machines Corporation
IEEE	Institute of Electrical and Electronics Engineers
IEEE-CIS	IEEE Computational Intelligence Society
k-NN	k-Nearest Neighbors
LightGBM	Light Gradient Boosting Machine
LIME	Local Interpretable Model-agnostic Explanations
LiteMORT	Lightweight Monotonic Optimal Regression Tree
LOF	Local Outlier Factor
LR	Logistic Regression
LSTM	Long Short-Term Memory
MCC	Matthews Correlation Coefficient
MDPI	Multidisciplinary Digital Publishing Institute
ML	Machine Learning
MLP	Multi-Layer Perceptron
PCA	Principal Component Analysis
PR	Precision-Recall
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RBF	Radial Basis Function
RF	Random Forest
RHSOFS	Rock Hyrax Swarm Optimization Feature Selection
RIPPER	Repeated Incremental Pruning to Produce Error Reduction
RQ	Research Question
RWN	Random Weight Network
SHAP	Shapley Additive Explanations
SMOTE	Synthetic Minority Over-sampling Technique
SLR	Systematic Literature Review
SL-SSNet	Sea Lion Optimization + Self-Supervised Network
SVM	Support Vector Machine
TabNet	Tabular Neural Network
TabTransformer	Transformer architecture for tabular data
TN	True Negative
TNR	True Negative Rate
TP	True Positive
TPC	Total Processing Cost
TPR	True Positive Rate
UCSD–FICO	University of California, San Diego—Fair Isaac Corporation
WoS	Web of Science
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

References

1. Hilal, W.; Gadsden, S.A.; Yawney, J. Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. *Expert Syst. Appl.* **2022**, *193*, 116429. [CrossRef]
2. Ali, A.; Abd Razak, S.; Othman, S.H.; Eisa, T.A.E.; Al-Dhaqm, A.; Nasser, M.; Elhassan, T.; Elshafie, H.; Saif, A. Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Appl. Sci.* **2022**, *12*, 9637. [CrossRef]
3. European Banking Authority (EBA); European Central Bank (ECB). 2024 Report on Payment Fraud. Available online: https://www.eba.europa.eu/sites/default/files/2024-08/465e3044-4773-4e9d-8ca8-b1cd031295fc/EBA_ECB%202024%20Report%20on%20Payment%20Fraud.pdf (accessed on 8 June 2025).
4. Nilson Report. Issue 1276; December 2024. Available online: https://nilsonreport.com/newsletters/1276/ (accessed on 13 June 2025).

5. Entrust. 2025 Identity Fraud Report. Available online: https://www.entrust.com/sites/default/files/documentation/reports/2025-identity-fraud-report.pdf (accessed on 13 June 2025).
6. Chaquet-Ulldemolins, J.; Moral-Rubio, S.; Muñoz-Romero, S. On the Black-Box Challenge for Fraud Detection Using Machine Learning (II): Nonlinear Analysis through Interpretable Autoencoders. *Appl. Sci.* **2022**, *12*, 3856. [CrossRef]
7. Baisholan, N.; Dietz, J.E.; Gnatyuk, S.; Turdalyuly, M.; Matson, E.T.; Baisholanova, K. FraudX AI: An Interpretable Machine Learning Framework for Credit Card Fraud Detection on Imbalanced Datasets. *Computers* **2025**, *14*, 120. [CrossRef]
8. Visbeek, S.; Acar, E.; den Hengst, F. Explainable Fraud Detection with Deep Symbolic Classification. In *Explainable Artificial Intelligence*; Longo, L., Lapuschkin, S., Seifert, C., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2024; Volume 2155, pp. 350–373. [CrossRef]
9. Kanika; Singla, J. Online Banking Fraud Detection System: A Review. *Int. J. Adv. Trends Comput. Sci. Eng.* **2019**, *8*, 959–962. [CrossRef]
10. Zhang, X.; Guo, F.; Chen, T.; Pan, L.; Beliakov, G.; Wu, J. A Brief Survey of Machine Learning and Deep Learning Techniques for E-Commerce Research. *J. Theor. Appl. Electron. Commer. Res.* **2023**, *18*, 2188–2216. [CrossRef]
11. Kaur, P.; Krishan, K.; Sharma, S.K.; Kanchan, T. ATM Card Cloning and Ethical Considerations. *Sci. Eng. Ethics* **2019**, *25*, 1311–1320. [CrossRef]
12. Onumadu, P.; Abroshan, H. Near-Field Communication (NFC) Cyber Threats and Mitigation Solutions in Payment Transactions: A Review. *Sensors* **2024**, *24*, 7423. [CrossRef]
13. du Preez, A.; Bhattacharya, S.; Beling, P.; Bowen, E. Fraud Detection in Healthcare Claims Using Machine Learning: A Systematic Review. *Artif. Intell. Med.* **2025**, *160*, 103061. [CrossRef]
14. Najar, A.V.; Alizamani, L.; Zarqi, M.; Hooshmand, E. A Global Scoping Review on the Patterns of Medical Fraud and Abuse: Integrating Data-Driven Detection, Prevention, and Legal Responses. *Arch. Public Health* **2025**, *83*, 43. [CrossRef]
15. Sulaiman, S.S.; Nadher, I.; Hameed, S.M. Credit Card Fraud Detection Challenges and Solutions: A Review. *Iraqi J. Sci.* **2024**, *65*, 2287–2303. [CrossRef]
16. Cherif, A.; Badhib, A.; Ammar, H.; Alshehri, S.; Kalkatawi, M.; Imine, A. Credit Card Fraud Detection in the Era of Disruptive Technologies: A Systematic Review. *J. King Saud Univ. Comput. Inf. Sci.* **2023**, *35*, 145–174. [CrossRef]
17. Alamri, M.; Ykhlef, M. Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques. *Electronics* **2022**, *11*, 4003. [CrossRef]
18. Kitchenham, B.; Charters, S. Guidelines for Performing Systematic Literature Reviews in Software Engineering, Version 2.3. 2007. Available online: https://www.researchgate.net/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering (accessed on 15 June 2025).
19. Hernandez Aros, L.; Bustamante Molano, L.X.; Gutiérrez-Portela, F.; Moreno Hernandez, J.J.; Rodríguez Barrero, M.S. Financial Fraud Detection through the Application of Machine Learning Techniques: A Literature Review. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 1130. [CrossRef]
20. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. [CrossRef]
21. Dal Pozzolo, A.; Boracchi, G.; Caelen, O.; Alippi, C.; Bontempi, G. Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 3784–3797. [CrossRef]
22. Worldline; Machine Learning Group–ULB. Credit Card Fraud Detection Dataset. Available online: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud (accessed on 7 August 2025).
23. Yang, Y.; Liu, C.; Liu, N. Credit Card Fraud Detection Based on CSat-Related AdaBoost. In Proceedings of the 8th International Conference on Computing and Pattern Recognition (ICCPR'19), Beijing, China, 23–25 October 2019; Association for Computing Machinery: New York, NY, USA, 2020; pp. 420–425. [CrossRef]
24. Taha, A.A.; Malebary, S.J. An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine. *IEEE Access* **2020**, *8*, 25579–25587. [CrossRef]
25. Sailusha, R.; Gnaneswar, V.; Ramesh, R.; Rao, G.R. Credit Card Fraud Detection Using Machine Learning. In Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 13–15 May 2020; pp. 1264–1270. [CrossRef]
26. Azhan, M.; Meraj, S. Credit Card Fraud Detection Using Machine Learning and Deep Learning Techniques. In Proceedings of the 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 3–5 December 2020; pp. 514–518. [CrossRef]
27. Sivanantham, S.; Dhinagar, S.R.; Kawin, P.; Amarnath, J. Hybrid Approach Using Machine Learning Techniques in Credit Card Fraud Detection. In *Advances in Smart System Technologies*; Suresh, P., Ed.; Advances in Intelligent Systems and Computing; Springer: Singapore, 2021; Volume 1163, pp. 243–251. [CrossRef]

28. Forough, J.; Momtazi, S. An Ensemble Deep Learning-Based Approach for Credit Card Fraud Detection. *Appl. Soft Comput.* **2021**, *99*, 106883. [CrossRef]
29. Trisanto, D.; Rismawati, N.; Mulya, M.F.; Kurniadi, F.I. Modified Focal Loss in Imbalanced XGBoost for Credit Card Fraud Detection. *Int. J. Intell. Eng. Syst.* **2021**, *14*, 350–358. [CrossRef]
30. Singh, A.K. Detection of Credit Card Fraud Using Machine Learning Algorithms. In Proceedings of the 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 9–10 December 2022; pp. 673–677. [CrossRef]
31. Gupta, Y. Using of Machine Learning Techniques to Detect Credit Card Frauds. In Proceedings of the 2022 OITS International Conference on Information Technology (OCIT), Bhubaneswar, India, 14–16 December 2022; pp. 124–128. [CrossRef]
32. Jain, V.; Kavitha, H.; Mohana Kumar, S. Credit Card Fraud Detection Web Application Using Streamlit and Machine Learning. In Proceedings of the 2022 IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India, 29–30 July 2022; pp. 1–5. [CrossRef]
33. Baker, M.R.; Mahmood, Z.N.; Shaker, E.H. Ensemble Learning with Supervised Machine Learning Models to Predict Credit Card Fraud Transactions. *RIA Rev. d'Intell. Artif.* **2022**, *36*, 509–518. [CrossRef]
34. Padhi, B.K.; Chakravarty, S.; Naik, B.; Pattanayak, R.M.; Das, H. RHSOFS: Feature Selection Using the Rock Hyrax Swarm Optimization Algorithm for Credit Card Fraud Detection System. *Sensors* **2022**, *22*, 9321. [CrossRef]
35. Sudarshana, K.; MylaraReddy, C.; Adhoni, Z.A. Classification of Credit Card Frauds Using Autoencoded Features. In *Intelligent Computing and Applications*; Rao, B.N.K., Balasubramanian, R., Wang, S.-J., Nayak, R., Eds.; Smart Innovation, Systems and Technologies; Springer: Singapore, 2023; Volume 315, pp. 9–17. [CrossRef]
36. Anagha, T.S.; Fathima, A.; Naik, A.D.; Goenka, C.; Devamane, S.B.; Thimmapurmath, A.R. Credit Card Fraud Detection Using Machine Learning Algorithms. In Proceedings of the 2023 International Conference on Computational Intelligence for Information, Security and Communication Applications (CIISCA), Bengaluru, India, 22–23 June 2023; pp. 419–424. [CrossRef]
37. Ndama, O.; En-Naimi, E.M. Credit Card Fraud Detection Using SVM, Decision Tree and Random Forest Supervised Machine Learning Algorithms. In Proceedings of the 6th International Conference on Big Data and Internet of Things (BDIoT 2022), Tangier, Morocco, 25–27 October 2022; Lecture Notes in Networks and Systems. Springer: Cham, Switzerland, 2023; Volume 625, pp. 316–327. [CrossRef]
38. Btoush, E.; Zhou, X.; Gururajan, R.; Chan, K.; Alsodi, O. Optimising Security: A Hybrid CNN-BiLSTM Model for Credit Card Cyber Fraud Detection. In Proceedings of the 2024 12th International Conference on Advanced Cloud and Big Data (CBD), Brisbane, Australia, 28 November–2 December 2024; pp. 380–385. [CrossRef]
39. Ramesh, S.; Simna, T.M.; Mohana. Analysis of Credit Card Fraudulent Transactions Using Machine Learning and Artificial Intelligence. In Proceedings of the 2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI), Coimbatore, India, 28–30 August 2024; pp. 1226–1231. [CrossRef]
40. Ileberi, E.; Sun, Y. A Hybrid Deep Learning Ensemble Model for Credit Card Fraud Detection. *IEEE Access* **2024**, *12*, 175829–175838. [CrossRef]
41. Juneja, S.; Bhati, B.S.; Goyal, R.; Atwal, S.; Maiti, S.; Chaudhry, N. Fraud Detection in Credit Card Using Machine Learning. In Proceedings of the Fifth Doctoral Symposium on Computational Intelligence (DoSCI 2024), Lucknow, India, 10 May 2024; Lecture Notes in Networks and Systems. Springer: Singapore, 2024; Volume 1085, pp. 1264–1270. [CrossRef]
42. Sruthi, S.; Emadaboina, S.; Jyotsna, C. Enhancing Credit Card Fraud Detection with Light Gradient-Boosting Machine: An Advanced Machine Learning Approach. In Proceedings of the 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS), Chikkaballapur, India, 18–19 April 2024; pp. 1–6. [CrossRef]
43. Rajesh, P.K.; Shreyanth, S.; Sarveshwaran, R.; Nithin Chary, V. Bayesian Optimized Random Forest Classifier for Improved Credit Card Fraud Detection: Overcoming Challenges and Limitations. In *Accelerating Discoveries in Data Science and Artificial Intelligence I (ICDSAI 2023)*; Lin, F.M., Patel, A., Kesswani, N., Sambana, B., Eds.; Springer Proceedings in Mathematics & Statistics; Springer: Cham, Switzerland, 2024; Volume 421, pp. 205–214. [CrossRef]
44. Manjula Devi, C.; Gobinath, A.; Padma Priya, S.; Adithiyaa, M.; Chandru, M.K.; Jothi, M. Next-Generation Anomaly Detection Framework Leveraging Artificial Intelligence for Proactive Credit Card Fraud Prevention and Risk Management. In Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 24–28 June 2024; pp. 1–6. [CrossRef]
45. Cherkaoui, R.; En-Naimi, E.M.; Kouissi, M. Combining Supervised and Unsupervised Machine Learning Methods for Improving Credit Card Fraud Detection. In Proceedings of the 7th International Conference on Networking, Intelligent Systems and Security (NISS '24), Meknes, Morocco, 18–19 April 2024; Association for Computing Machinery: New York, NY, USA, 2024; pp. 1–5, Article 47. [CrossRef]
46. Renganathan, K.K.; Karuppiah, J.; Pathinathan, M.; Raghuraman, S. Credit card fraud detection with advanced graph-based machine learning techniques. *Indones. J. Electr. Eng. Comput. Sci.* **2024**, *35*, 1963–1975. [CrossRef]

47. Planinić, D.; Popović-Bugarin, V. Credit Card Fraud Detection Using Supervised Learning Algorithms. In Proceedings of the 2024 28th International Conference on Information Technology (IT), Žabljak, Montenegro, 21–24 February 2024; pp. 1–4. [CrossRef]
48. Rawashdeh, E.; Al-Ramahi, N.; Ahmad, H.; Zaghloul, R. Efficient credit card fraud detection using evolutionary hybrid feature selection and random weight networks. *Int. J. Data Netw. Sci.* **2024**, *8*, 463–472. [CrossRef]
49. Conteh, A.A.; Zhou, J. Credit Card Fraud Detection with Imbalanced Small Data Using TabTransformer and Cost-Sensitive Learning. In *Cyberspace Simulation and Evaluation (CSE 2024)*; Xu, G., Zhou, W., Zhang, J., Zhang, Y., Jia, Y., Eds.; Communications in Computer and Information Science; Springer: Singapore, 2025; Volume 2422, pp. 35–50. [CrossRef]
50. Ram, M.A.; Geetha, S.; Vignesh, U. Machine Learning Techniques for Credit Card Fraudulent Transaction Detection. In *Fifth International Conference on Computing and Network Communications (CoCoNet 2023)*; Thampi, S., Siarry, P., Atiquzzaman, M., Trajkovic, L., Lloret Mauri, J., Eds.; Lecture Notes in Electrical Engineering; Springer: Singapore, 2025; Volume 1219, pp. 122–125. [CrossRef]
51. Kasula, V.K.; Yenugula, M.; Yadulla, A.R.; Konda, B.; Ayyamgari, S. An improved machine learning technique for credit card fraud detection. *Edelweiss Appl. Sci. Technol.* **2025**, *9*, 3093–3109. [CrossRef]
52. Gao, Y.; Zhang, S.; Lu, J. Machine Learning for Credit Card Fraud Detection. In Proceedings of the 2021 International Conference on Computational Intelligence and Research (ICCIR), Guangzhou, China, 18–20 June 2021; Association for Computing Machinery: Guangzhou, China, 2021; pp. 213–219. [CrossRef]
53. Howard, A.; Bouchon-Meunier, B.; IEEE CIS; Inversion; Lei, J.; Lynn@Vesta; Marcus2010; Abbass, H. IEEE-CIS Fraud Detection. Available online: https://kaggle.com/competitions/ieee-fraud-detection (accessed on 10 August 2025).
54. Liu, X.; Yan, K.; Kara, L.B.; Nie, Z. CCFD-Net: A Novel Deep Learning Model for Credit Card Fraud Detection. In Proceedings of the 2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI), Las Vegas, NV, USA, 10–12 August 2021; IEEE: Los Angeles, CA, USA, 2021; pp. 9–16. [CrossRef]
55. Bakhtiari, S.; Nasiri, Z.; Vahidi, J. Credit card fraud detection using ensemble data mining methods. *Multimed. Tools Appl.* **2023**, *82*, 29057–29075. [CrossRef]
56. Shenoy, K. Credit Card Transactions Fraud Detection Dataset (Sparkov Simulation). Available online: https://www.kaggle.com/datasets/kartik2112/fraud-detection (accessed on 10 August 2025).
57. Kim, K.C.L.; Mustapha, A.; Palaniappan, V.; Mun, W.K.; Kasinathan, V. A Comparative Analysis of Credit Card Detection Models. In Proceedings of the 8th International Conference on the Applications of Science and Mathematics (EduTA 2022), Malaysia, 17–19 October 2023; Springer Proceedings in Physics. Springer: Singapore, 2023; Volume 294. [CrossRef]
58. Jain, S.; Sharma, N.; Kumar, M. FraudFort: Harnessing Machine Learning for Credit Card Fraud Detection. In Proceedings of the 2024 First International Conference on Technological Innovations and Advance Computing (TIACOMP), Bali, Indonesia, 29–30 June 2024; pp. 41–46. [CrossRef]
59. Harish, S.; Lakhanpal, C.; Jafari, A.H. Leveraging graph-based learning for credit card fraud detection: A comparative study of classical, deep learning and graph-based approaches. *Neural Comput. Appl.* **2024**, *36*, 21873–21883. [CrossRef]
60. Ghahfarokhi, A.F.; Mansouri, T.; Sadeghi Moghaddam, M.R.; Bahrambeik, N.; Yavari, R.; Sani, M.F. Credit card fraud detection using asexual reproduction optimization. *Kybernetes* **2022**, *51*, 2852–2876. [CrossRef]
61. Singh, A.; Jain, A. An Efficient Credit Card Fraud Detection Approach Using Cost-Sensitive Weak Learner with Imbalanced Dataset. *Comput. Intell.* **2022**, *38*, 2035–2055. [CrossRef]
62. Lopez-Rojas, E. BankSim: Synthetic Data from a Financial Payment System. Available online: https://www.kaggle.com/datasets/ealaxi/banksim1 (accessed on 10 August 2025).
63. Lopez-Rojas, E.A.; Axelsson, S. Banksim: A Bank Payments Simulator for Fraud Detection Research. In Proceedings of the 26th European Modeling and Simulation Symposium (EMSS 2014), Bordeaux, France, 10–12 September 2014; DIME University of Genoa: Genoa, Italy; pp. 144–152, ISBN 978-88-97999-32-4.
64. Prusti, D.; Das, D.; Rath, S.K. Credit Card Fraud Detection Technique by Applying Graph Database Model. *Arab. J. Sci. Eng.* **2021**, *46*, 1–20. [CrossRef]
65. Prusti, D.; Behera, R.K.; Rath, S.K. Hybridizing Graph-Based Gaussian Mixture Model with Machine Learning for Classification of Fraudulent Transactions. *Comput. Intell.* **2022**, *38*, 2134–2160. [CrossRef]
66. Lopez-Rojas, E. PaySim Dataset: Synthetic Financial Datasets for Fraud Detection. Available online: https://www.kaggle.com/datasets/ealaxi/paysim1 (accessed on 10 August 2025).
67. Lopez-Rojas, E.A.; Elmir, A.; Axelsson, S. PaySim: A Financial Mobile Money Simulator for Fraud Detection. In Proceedings of the 28th European Modeling and Simulation Symposium (EMSS 2016), Larnaca, Cyprus, 26–28 September 2016.
68. Khang, V.H.; Anh, C.T.; Thuan, N.D. Detecting Fraud Transaction Using RIPPER Algorithm Combined with Ensemble Learning Model. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 336–345. [CrossRef]
69. Atia, H.A.; Aboul-Ela, M.; Reyad, C.A.; Awad, N.A. Online Payments Fraud Detection Using Machine Learning Techniques. In Proceedings of the 2024 Intelligent Methods, Systems, and Applications (IMSA), Giza, Egypt, 3–5 June 2024; pp. 402–409. [CrossRef]

70. IBM. TabFormer (2021). Available online: https://github.com/IBM/TabFormer (accessed on 10 August 2025).
71. Choudhury, N.R. Credit Card Fraud Data. Available online: https://www.kaggle.com/datasets/neharoychoudhury/credit-card-fraud-data/data (accessed on 10 August 2025).
72. Karthik, M.; Saigeeta, S.B.; Reddy, P.K.S.C.; Reddy, P.R.D.; Srinivas, M. Integrating Machine Learning and Deep Learning for Accurate Fraud Detection. In Proceedings of the 2025 International Conference on Artificial Intelligence and Data Engineering (AIDE), Nitte, India, 23–25 January 2025; pp. 565–569. [CrossRef]
73. Karnavou, E.; Cascavilla, G.; Marcelino, G.; Geradts, Z. I Know You're a Fraud: Uncovering Illicit Activity in a Greek Bank's Transactions with Unsupervised Learning. *Expert Syst. Appl.* **2025**, *288*, 128148. [CrossRef]
74. Saranya, N.; Devi, M.K.; Mythili, A.; H, S.P. Data Science and Machine Learning Methods for Detecting Credit Card Fraud. *Sci. Temper* **2023**, *14*, 840–844. [CrossRef]
75. Ndama, O.; Bensassi, I.; En-Naimi, E.M. Optimizing Credit Card Fraud Detection: A Deep Learning Approach to Imbalanced Datasets. *Int. J. Electr. Comput. Eng.* **2024**, *14*, 4802–4814. [CrossRef]
76. Omair, B.; Alturki, A. A Systematic Literature Review of Fraud Detection Metrics in Business Processes. *IEEE Access* **2020**, *8*, 26893–26903. [CrossRef]
77. Fanai, H.; Abbasimehr, H. A Novel Combined Approach Based on Deep Autoencoder and Deep Classifiers for Credit Card Fraud Detection. *Expert Syst. Appl.* **2023**, *217*, 119562. [CrossRef]
78. Baisholan, N.; Baisholanova, K.; Kubayev, K.; Alimzhanova, Z.; Baimuldina, N. Corporate Network Anomaly Detection Methodology Utilizing Machine Learning Algorithms. *Smart Sci.* **2024**, *12*, 666–678. [CrossRef]
79. Ullah, I.; Rios, A.; Gala, V.; McKeever, S. Explaining Deep Learning Models for Tabular Data Using Layer-Wise Relevance Propagation. *Appl. Sci.* **2022**, *12*, 136. [CrossRef]
80. Ikermane, M.; Mohy-eddine, M.; Rachidi, Y. Credit Card Fraud Detection: Comparing Random Forest and XGBoost Models with Explainable AI Interpretations. In *Innovative Technologies on Electrical Power Systems for Smart Cities Infrastructure. ICESST 2024*; Aboudrar, I., Ilahi Bakhsh, F., Nayyar, A., Ouachtouk, I., Eds.; Sustainable Civil Infrastructures; Springer: Cham, Switzerland, 2025. [CrossRef]
81. Chavakula, S.J.; Albert, C.A.J.; Ebenezer, E.; Bhagat, M.H.; Mahamuni, C.V. Explainable AI (XAI) Using SHAP and LIME for Financial Fraud Detection and Credit Scoring. In Proceedings of the 2025 International Conference on Advanced Computing Technologies (ICoACT), Sivalasi, India, 14–15 March 2025; IEEE: New York, NY, USA, 2025; pp. 1–9. [CrossRef]
82. Raufi, B.; Finnegan, C.; Longo, L. A Comparative Analysis of SHAP, LIME, Anchors, and DiCE for Interpreting a Dense Neural Network in Credit Card Fraud Detection. In *xAI 2024: International Conference on eXplainable AI*; Springer: Cham, Switzerland, 2024; pp. 365–383. [CrossRef]
83. Adams, J.M.; Reinert, G.; Szpruch, L.; Maple, C.; Elliott, A. Individualised Counterfactual Examples Using Conformal Prediction Intervals. *Proc. Mach. Learn. Res.* **2025**, *266*, 425–444.
84. Achituve, I.; Kraus, S.; Goldberger, J. Interpretable Online Banking Fraud Detection Based on Hierarchical Attention Mechanism. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 1–9. [CrossRef]
85. Al-Daoud, K.I.; Abu-AlSondos, I.A. Robust AI for Financial Fraud Detection in the GCC: A Hybrid Framework for Imbalance, Drift, and Adversarial Threats. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 121. [CrossRef]
86. Gadi, M.F.A.; Wang, X.; do Lago, A.P. Credit Card Fraud Detection with Artificial Immune System. In *Artificial Immune Systems. ICARIS 2008*; Bentley, P.J., Lee, D., Jung, S., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5132, pp. 119–131. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.