Ref	Research Objective	Paper Type	Fraud Detection Focus	Main Conclusions	Algorithm Type(s)			Algorithm Details	Data Source Description	Preprocessing & Feature Engineering Details	Evaluation Metrics	Problem(s) addressed
					Super-vised	Unsuper-vised	Semi					
(Li et al., 2021)	To address class imbalance with overlap in credit card fraud detection using a hybrid machine-learning approach.	Technical Paper	Credit card fraud detection	The proposed method outperformed traditional approaches, improved efficiency, minimized information loss, and reduced time consumption for large dataset processing. DWE facilitated effective hyper-parameter selection.	X	X	X	- Anomaly detection models (e.g., Isolation Forest, One-Class SVM, Auto-Encoder) - Non-linear classifiers (e.g., Random Forest, ANN)	- Fraud detection dataset of European cardholders from Kaggle - Large, real electronic transaction dataset from a financial company in China	- Dynamic Weighted Entropy (DWE) for dataset quality evaluation and class balance - Anomaly detection on minority samples for segregation of overlapped transactions	- Accuracy - AUC PR - F1-Score - Precision	- Class Imbalance - Class Overlapping
(Lacruz & Saniie, 2021)	To explore the application of AI and machine learning algorithms, specifically Autoencoders and Logistic Regression, in detecting credit card fraud, aiming to achieve high accuracy in fraud transaction predictions.	Technical Paper	Credit card fraud detection	Both Autoencoders and Logistic Regression showed high accuracy in detecting fraud transactions, with promising results demonstrating the efficacy of AI in improving financial transaction security.	X	X		- Autoencoders - Logistic Regression (LR)	- Fraud detection dataset of European cardholders from Kaggle	- Min/Max scaling to normalize features - Minority Over-sampling Technique (SMOTE) to address class imbalance - Correlation Analysis to identify attributes that heavily influence the prediction of fraud - Outliers are removed to improve model accuracy	- Accuracy - F1-Score - Precision - Recall	- Class Imbalance - High Dimensionality
(Stojanović et al., 2021)	To develop and evaluate machine learning methods for detecting fraudulent transactions in the Fintech domain. The study focuses on using anomaly detection methods and evaluating their effectiveness in identifying suspicious activities in financial datasets.	Technical Paper	Transaction fraud detection	ML methods, particularly ensemble approaches, significantly contribute to fraud detection in financial transactions. The proposed methods demonstrate varying degrees of success, with ensemble methods generally outperforming outlier detection methods, especially in synthetic datasets.	X	X		- Local Outlier Factor (LOF) - Isolation Forest (IF) - Elliptic Envelope (eEnvelop) - Random Forest (RF) - Adaptive Boosting (AdaBoost) - Extreme Gradient Boosting (XGBoost) - Neural Networks	- Fraud detection dataset of European cardholders from Kaggle - Synthetic Financial datasets (PaySim and BankSim)	- Data Statistics and Visualization to gain an initial overview - Categorical Variable Encoding to convert categorical data into numerical - Box–Cox transformation to skewed features to approximate normal distribution - Additional Feature Engineering to create new features	- ROC - True Negative Rate (TNR) - True Positive Rate (TPR) - Validation accuracy	- Class Imbalance - Concept Drift - Real-Time Processing
(Islam et al., 2023)	To address the challenge of anomaly detection in credit card transactions, which is complicated by imbalanced datasets and overlapping class samples.	Technical Paper	Credit card fraud detection	The CCAD model outperforms existing models in detecting anomalies, particularly from minority class instances, by effectively managing the issues of class imbalance and model overfitting.	X	X		Base learner - Elliptic Envelope (eEnvelope) - Isolation Forest (IF) - Local Outlier Factor (LOF) - One-Class SVM (OCSVM) Meta-learner - XGBoost	- Credit Card Fraud (CCF) dataset - Credit Card Default Payment (CCDP) dataset	- Integrity Check to ensure no missing or null values in the dataset - Removing Duplicates to eliminate redundant information - XGBoost Feature Importance - Stratified Sampling to handle the imbalanced dataset - K-fold Cross-Validation to prevent Overfitting	- Accuracy - AUC - F1-Score - Precision - Recall	- Class Imbalance - Class Overlapping - Overfitting
(Murugan et al., 2023)	To improve the accuracy and detection rate of credit card fraud identification by using Support Vector Machine	Technical Paper	Credit card fraud detection	The proposed SVM-IG method showed better performance in detection rates, precision, and accuracy compared to Bayesian	X			- Support Vector Machine with Information Gain (SVM-IG)	- Taiwan credit dataset - German credit dataset	- Discretization to convert continuous features to discrete intervals using entropy	- Accuracy - Precision - Pruning time	- Class Imbalance - High Dimensionality

	(SVM) and enhancing performance through effective feature selection and normalization techniques.			Network Classifier (BNC), CART-based Random Forest (CARTRF), and Random-Tree-Based Random Forest (RTBRF) methods.				- CART-based Random Forest (CARTRF) - Random-Tree-Based Random Forest (RTBRF)		- Min-Max Normalization to scale features to a specific range to improve training efficiency - Information Gain (IG) to select the most relevant attributes by calculating information gain - Apriori Algorithm to identify and prune frequent item sets to enhance performance		
(Z. Zhang et al., 2020)	To develop a fraud detection method specifically for low-frequency users whose transaction volumes are low. The aim is to improve the accuracy of fraud detection for these users by constructing individual behavior models that incorporate group behavior and current transaction status.	Technical Paper	Transaction fraud detection	The proposed method significantly improves the detection accuracy for low-frequency users compared to existing models. The method combines individual transaction history, group behavior extracted through DBSCAN clustering, and current transaction status using a sliding window mechanism.	X			- Naïve Bayes	- Dataset from a domestic bank	- Extract user behavior benchmarks from historical transactions - Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for clustering to identify group behaviors from normal and fraudulent transactions - Sliding window mechanism to assess the current transaction status	- Disturbance Rate - F1-Score - Precision - Recall	- Inaccurate Fraud Detection for Low-Frequency Users - Dynamic User Behavior
(Kalid et al., 2020)	To improve anomaly detection in credit card transactions using a Multiple Classifiers System (MCS) specifically designed to handle datasets with unbalanced class distribution and overlapping class samples.	Technical Paper	Credit card fraud detection	The MCS approach significantly improved the detection of minority anomalies in credit card datasets compared to traditional single classifiers or previous research methods.	X			- Multiple Classifiers System (MCS)	- Worldline and ULB Machine Learning Group datasets - Payment data of credit card holders from a Taiwan bank	- Employed strategies to manage the unbalanced class distribution in both datasets - Utilized visualization to understand and address the issue of overlapping class samples - Visualized pairwise relationships of attributes to identify overlapping and informative features	- Accuracy - True Negative Rate (TNR) - True Positive Rate (TPR)	- Class Imbalance - Class Overlapping
(Esenogho et al., 2022)	To develop a robust solution for credit card fraud detection using an LSTM neural network ensemble and adaptive boosting with hybrid data resampling for handling imbalanced datasets.	Technical Paper	Credit card fraud detection	The proposed LSTM ensemble outperformed conventional algorithms, achieving high sensitivity (0.996) and specificity (0.998) and demonstrating the effectiveness of feature engineering via resampling.	X			- Long Short-Term Memory (LSTM) neural networks Ensemble with - Adaptive Boosting (AdaBoost)	- Machine Learning Group of Université Libre de Bruxelles (ULB) dataset	- SMOTE-ENN to balance the dataset by oversampling the minority class and undersampling the majority class to remove overlapping instances	- AUC - Recall - True Negative Rate (TNR)	- Class Imbalance - Class Overlapping
(Dileep et al., 2021)	To improve the detection of credit card fraud using machine learning algorithms, specifically Decision Trees and Random Forests, and to evaluate the effectiveness of these models using real-world data.	Technical Paper	Credit card fraud detection	Both algorithms demonstrated good precision levels in identifying fraudulent transactions, with Random Forest generally providing better results due to its ensemble approach	X			- Decision Trees - Random Forest	- Credit card data from a financial institution	- Clustering Model to categorize transactions into clusters to identify patterns - Gaussian Mixture Technique to model probability density to detect anomalies -Bayesian Network Technique to construct a DAG to identify dependencies and potential fraud areas	- MCC	- Class Imbalance - Scalability

(Mniai et al., 2023)	To develop an effective framework for detecting fraudulent credit card transactions, addressing the challenges posed by high-dimensional and imbalanced datasets.	Technical Paper	Credit card fraud detection	The proposed fraud detection framework outperformed traditional methods, improving model accuracy from 90% to 93% by addressing data imbalance and optimizing feature selection. The optimized SVDD model and PSLPSO algorithm demonstrated the importance of relevant features for better prediction performance.	X			- Support Vector Data Description (SVDD) with - Polynomial Self Learning PSO (PSLPSO)	- Fraud detection dataset of European cardholders from Kaggle	- Normalization and Scaling to ensure all input values are within a predefined interval - Undersampling to reduce the majority class data points to match the minority class using the imblearn technique - SelectKBest to rank features based on their importance - Recursive Feature Elimination (RFE) to eliminate the least significant features	- Accuracy - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance - High Dimensionality
(Taha & Malebary, 2020)	To develop an intelligent approach for detecting fraud in credit card transactions using an optimized Light Gradient Boosting Machine (OLightGBM).	Technical Paper	Credit card fraud detection	The OLightGBM approach significantly outperforms other methods, demonstrating high efficiency and effectiveness in detecting credit card fraud.	X			- OLightGBM (Optimized Light Gradient Boosting Machine)	- Credit Card Transactions (European Transactions) dataset - UCSD-FICO Data Mining Contest 2009 dataset	- Cross Validation to ensure accurate model evaluation - Information Gain (IG) used by LightGBM to select significant features, reducing data dimensionality	- Accuracy - AUC - F1-Score - Precision - Recall	- Class Imbalance - High Dimensionality - Overfitting
(Seera et al., 2021)	To investigate the effectiveness of using aggregated features for payment card fraud detection by employing a variety of statistical and machine learning models on both publicly available and real transaction records.	Technical Paper	Payment card fraud detection	Aggregated features provided significant improvements in fraud detection over original features. The use of real transaction data added practical relevance to the findings, demonstrating the robustness of the models in actual financial environments.	X			- Thirteen statistical and machine learning methods were employed - Including SVMs (Support Vector Machines), ANNs (Artificial Neural Networks), Decision Trees, and deep learning models	- Publicly available dataset - Payment card transaction database from a financial institution in Malaysia	- Feature aggregation - Genetic algorithm to optimize feature selection process	- Accuracy - AUC - MCC	- Class Imbalance - Concept drift
(Reddy et al., 2023)	To develop an advanced financial fraud detection system that leverages data preprocessing, Kernel PCA for feature extraction, and a hybrid CNN-BiLSTM neural network to enhance detection accuracy and efficiency.	Technical Paper	Transaction fraud detection	The proposed CNN-BiLSTM hybrid model significantly improves financial fraud detection accuracy, achieving approximately 97.5% accuracy. This model effectively combines the strengths of Convolutional Neural Networks (CNN) for extracting nuanced features and Bidirectional Long Short-Term Memory (BiLSTM) networks.	X			- CNN-BiLSTM hybrid model (combination Convolutional Neural Networks and Bidirectional Long Short-Term Memory)	- Not specifically mentioned	- Data Cleaning - Min-Max normalization to transform the data to a uniform range - Kernel Principal Component Analysis (KPCA) to transform the data into a higher-dimensional space	- Accuracy - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance - High Dimensionality
(Alkhatib et al., 2021)	To develop a highly efficient model for credit card fraud detection using a deep learning approach, specifically targeting the dataset provided by the IEEE-CIS in cooperation with Vesta Company. The aim is to achieve high accuracy and	Technical Paper	Credit card fraud detection	The proposed deep learning model, which utilizes a deep neural network with seven dense layers, significantly outperforms previous models in detecting credit card fraud. The model achieved a high score of 99.1% in the area under the ROC curve, indicating its effectiveness in distinguishing	X			- Deep Neural Network (DNN)	- IEEE-CIS Fraud Detection dataset	- Data Cleaning to remove missing values - Standard Scaler and Label Encoder methods to transform features - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem	- Accuracy - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance - Missing values

	robustness in identifying fraudulent transactions.			between fraudulent and legitimate transactions.							
(Jiang & Liao, 2023)	To develop a credit card fraud detection method that leverages DeepInsight and deep learning to convert non-image transaction data into structured images, which are then classified using a parallel convolutional neural network (CNN) to enhance detection accuracy and performance metrics.	Technical Paper	Credit card fraud detection	The proposed method, combining DeepInsight and a parallel CNN model, outperforms traditional methods in detecting credit card fraud. It achieves superior results in terms of accuracy, true positive rate, true negative rate, and Matthews correlation coefficient, especially when addressing data imbalance through ADASYN sampling.	X			- Parallel Convolutional Neural Network (CNN)	- Fraud detection dataset of European cardholders from Kaggle	- Adaptive Synthetic Sampling (ADASYN) to address the data imbalance problem - Min-max normalization to scale the data for consistency - Data Transformation with DeepInsight to convert non-image credit card transaction data into structured 8x8 images	- Accuracy - MCC - True Negative Rate (TNR) - True Positive Rate (TPR)	- Class Imbalance
(Karthika & Senthilselvi, 2022)	To develop an ensemble-based machine learning model for credit card fraud detection. The goal is to enhance the prediction accuracy and efficiency by addressing issues related to data imbalance and feature selection, leveraging multiple machine learning techniques to identify fraudulent transactions effectively.	Technical Paper	Credit card fraud detection	The proposed ensemble-based approach, which combines RFE, SMOTE, and multiple machine learning classifiers, significantly improves the accuracy and efficiency of credit card fraud detection. The ETC performed the best among the evaluated classifiers, achieving high accuracy and F1-score.	X			- Passive Aggressive Classifier (PAC) - Linear Discriminant Analysis (LDA) - Radius Neighbors Classifier (RNC) - Bernoulli Naïve Bayesian (BNB) - Gaussian Naïve Bayesian (GNB) - Extra Tree Classifier (ETC)	- UCSD-FICO dataset - CCF dataset	- Dataset Merging - Removing missing values - Standard Scaler and Label Encoder methods to transform features - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem - Recursive Feature Elimination (RFE) to identify and select the most predictive features	- Accuracy - F1-Score - Precision - Recall	- Class Imbalance
(Karthik et al., 2022)	To develop a novel credit card fraud detection model that combines ensemble learning techniques, specifically boosting and bagging, to improve the detection of fraudulent transactions. The goal is to address issues related to data imbalance and enhance the detection accuracy and efficiency of fraud detection systems.	Technical Paper	Credit card fraud detection	The proposed hybrid ensemble model, which integrates boosting and bagging techniques, outperforms state-of-the-art fraud detection methods. It demonstrates superior performance in terms of true positive rate, false positive rate, true negative rate, false negative rate, detection rate, accuracy, and area under the curve (AUC)	X			- AdaBoost - Random Forest (RF) - Extra Trees Classifier (Extremely Randomized Trees)	- Brazilian bank dataset - UCSD-FICO dataset	- Data Distribution Analysis to understand the distribution - Outlier Detection to identify and remove outliers - Noise Elimination to improve data quality - Adaboost for feature selection	- Accuracy - AUC - Detection Rate - False Negative Rate (FNR) - False Positive Rate (FPR) - True Negative Rate (TNR) - MCC - Precision - Recall	- Class Imbalance - Dynamic Fraudulent Behavior - Noisy data - Scalability
(Mijwil & Salem, 2020)	To apply and compare three machine learning classifiers - Naïve Bayes, C4.5 decision trees, and Bagging Ensemble Learner - for credit card fraud detection, evaluating their performance to identify the most effective model.	Technical Paper	Credit card fraud detection	The C4.5 Decision Tree classifier exhibited the highest precision among the tested classifiers, indicating its robustness in identifying fraudulent transactions more accurately than the other models.	X			- Naïve Bayes - C4.5 Decision Trees - Bagging Ensemble	- Fraud detection dataset of European cardholders from Kaggle	- Not specifically mentioned	- Precision - Recall - PRC	- Class Imbalance
(Jabeen et al., 2023)	To develop an effective machine learning-based credit card fraud detection system.	Technical Paper	Credit card fraud detection	The use of SMOTE significantly improves the performance of all evaluated	X			- Decision Trees (DT) - Logistic	- Fraud detection dataset of European	- Data Splitting - SMOTE (Synthetic Minority Over-sampling	- Accuracy - AUC - F1-Score	- Class Imbalance

	The study aims to identify and predict fraudulent transactions by addressing the issue of data imbalance in the dataset using the Synthetic Minority Oversampling Technique (SMOTE).			machine learning models in terms of all considered metrics.				Regression (LR) - Random Forests (RF)	cardholders from Kaggle	Technique) to address the imbalance problem	- MCC - Precision - Recall	
(Vengatesan et al., 2020)	To compare the accuracy of different machine learning techniques, specifically logistical regression and K-Nearest Neighbors, to identify the most effective method for predicting fraudulent transactions.	Technical Paper	Credit card fraud detection	KNN algorithm performs better in detecting credit card fraud compared to the logistical regression algorithm. KNN achieved a precision value of 0.95, a recall value of 0.72, and an F1-score of 0.82, indicating its higher effectiveness in fraud detection.	X			- Logistic Regression (LR) - K-Nearest Neighbors (KNN)	- Not explicitly mentioned	- Data Cleaning - Error, Noise, and Inconsistency Handling to clean dataset from errors, noise, or inconsistent values	- Accuracy - F1-Score - Precision - Recall	- Class Imbalance - Noisy data
(Kaur et al., 2022)	To develop an efficient system for detecting fraudulent credit card transactions using ML and DL techniques. The study aims to train various models on highly imbalanced transactional data, compare their accuracies, and identify the most suitable model for fraud detection.	Technical Paper	Credit card fraud detection	The ensemble of a deep learning model and the K-Nearest Neighbors (KNN) model trained on oversampled data provides the best overall performance in detecting fraudulent transactions.	X			- Ensemble of the K-Nearest Neighbors (KNN) and - DL models	- Fraud detection dataset of European cardholders from Kaggle	- Standardization to eliminate the mean and scale to unit variance - Removal of irrelevant features - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem	- Accuracy - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance - High Feature Dimensionality
(Pranavi et al., 2022)	To compare the performance of various machine learning algorithms (Decision Tree, K-Nearest Neighbors, Logistic Regression, and Random Forest) and improve fraud detection accuracy, especially reducing false negatives, by applying oversampling techniques such as SMOTE.	Technical Paper	Credit card fraud detection	The Random Forest algorithm outperforms other classifiers in terms of precision, accuracy, recall, F1-score, MCC, and F1-score. When combined with the SMOTE to address data imbalance, the Random Forest algorithm achieved 100% recall, effectively detecting all fraudulent transactions while minimizing false negatives.	X			- Decision Tree (DT) - K-Nearest Neighbors (KNN) - Logistic Regression (LR) - Random Forest (RF)	- ULB Machine Learning Group dataset	- Handling Missing Values - Dataset Splitting - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem	- Accuracy - F1-Score - MCC - Precision - Recall - False Positive (FP) - False Negative (FN) - True Negative (TN) - True Positive (TP)	- Class Imbalance
(Alarfaj et al., 2022)	To address issues such as high-class imbalance in the data, changes in fraud patterns, and high rates of false alarms. The research focuses on improving detection accuracy and minimizing false negatives by leveraging state-of-the-art DL techniques, particularly CNNs.	Technical Paper	Credit card fraud detection	The proposed model, which involves various architectures of CNNs, achieved superior results with optimized accuracy (99.9%), F1-score (85.71%), precision (93%), and area under the curve (AUC) (98%).	X			- Convolutional neural networks (CNN) - ML algorithms like Extreme Learning Machine (ELM), Decision Tree, K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), Logistic	- Fraud detection dataset of European cardholders from Kaggle	- Principal Component Analysis (PCA) to reduce the dataset's dimensionality - Data Cleaning to remove any inconsistencies or errors - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem - PCA Transformation - StandardScaler to standardize features	- Accuracy - AUC - F1-Score - PRC - Precision - Recall	- Class Imbalance - Dynamic Fraudulent Behavior - High False Positives - High Feature Dimensionality

								Regression, and XG Boost			
(Alshammari et al., 2022)	To develop an efficient credit card fraud detection system using Big Data analytics, specifically leveraging Apache Spark and various machine learning models to enhance classification performance.	Technical Paper	Credit card fraud detection	The integration of Apache Spark with ML models significantly improves the classification accuracy of credit card fraud detection systems. The Gradient Boosting model demonstrated the best overall performance, and a high recall value.	X			- Various ML models like Logistic Regression (LR), Gradient Boosting (GB), Random Forest (RF), and Support Vector Machine (SVM)	- Fraud detection dataset of European cardholders from Kaggle	- Visualization techniques to understand the distribution and the correlation between different features - Analyzing the distribution of fraud versus non-fraud transactions	- Accuracy - Precision - Recall	- Class Imbalance - Scalability
(Kumar et al., 2023)	To achieve high precision, high detection rates of fraudulent activities, and a low number of false positives by analyzing customer behaviors and transaction patterns.	Technical Paper	Credit card fraud detection	The Random Forest algorithm outperforms other machine learning techniques in detecting fraudulent credit card transactions. The algorithm achieved the highest accuracy (94.4%) and Matthews Correlation Coefficient (MCC) scores.	X			- Decision Trees (DT) - Logistic Regression (LR) - Random Forest (RF) - Naive Bayes - Neural Networks - Support Vector Machines (SVM)	- Fraud detection dataset of European cardholders from Kaggle	- Standardization - Customer behavior patterns analyzed for deviations - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem	- Accuracy - Confusion Matrix - F1-Score - MCC - Precision - Recall - ROC AUC	- Class Imbalance - Dynamic User Behavior
(El Naby et al., 2021)	To investigate the effectiveness of Multi-Layer Perceptron (MLP) and Convolutional Neural Network (CNN) models on a credit card fraud detection dataset, applying SMOTE for data balancing to improve model accuracy.	Technical Paper	Credit card fraud detection	The proposed OSCNN (Over Sampling with Convolution Neural Network) model, which combines oversampling (SMOTE) with CNN, achieved superior performance with an accuracy of 98%, surpassing other models including MLP and MLP with SMOTE.	X			- Multi-layer Perceptron (MLP) - Convolutional Neural Network (CNN)	- Fraud detection dataset of European cardholders from Kaggle	- Dataset Division - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem	- Accuracy - Confusion Matrix - Precision - Recall	- Class Imbalance - Overfitting
(Forough & Momtazi, 2021)	To develop an ensemble model for credit card fraud detection using deep recurrent neural networks and a novel voting mechanism based on artificial neural networks.	Technical Paper	Credit card fraud detection	The proposed ensemble model, which uses either Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) networks as base classifiers and aggregates their outputs using a Feed-Forward Neural Network (FFNN), significantly outperforms state-of-the-art models.	X			- Long Short-Term Memory (LSTM) - Gated Recurrent Units (GRU) - Feedforward Neural Network (FFNN)	- Fraud detection dataset of European cardholders from Kaggle - Datasets from a Brazilian bank	- Data Splitting - Standardization	- AUC PR - Precision - Recall - ROC AUC - F1-Score	- Class Imbalance - Concept Drift - Real-Time Processing
(Isangediok & Gajamannage, 2022)	To investigate the effectiveness of various machine learning (ML) techniques in detecting fraud, particularly focusing on handling class imbalance in fraud detection datasets.	Technical Paper	Credit card fraud detection	The XGBoost model trained on the original imbalanced data consistently outperforms other classifiers and resampling techniques in terms of AUC ROC and AUC PR for both phishing website URLs and credit card fraud datasets.	X			- Logistic regression (LR) - Decision Trees (DT) - Random Forest (RF) - Extreme Gradient Boosting (XGBoost)	- ULB Machine Learning Group dataset - Phishing Website URLs dataset	- Z-scoring to improve model training efficiency - Random Under Sampler (RUS) to balance class distribution - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem - SMOTE Edited Nearest Neighbor (SMOTEENN) to oversample	- AUC PR - False Positives (FP) - False Negatives (FN) - Precision - Recall - ROC AUC	- Class Imbalance

(AbdulSattar & Hammad, 2020)	To classify credit card transactions as either fraudulent or legitimate using five different machine learning algorithms and to evaluate and compare their performance.	Technical Paper	Credit card fraud detection	Random Forest (RF) consistently showed superior performance based on kappa statistics and MCC values across both datasets. All classifiers achieved high accuracy, but SGD showed the weakest performance in comparison.	X			- Stochastic Gradient Descent (SGD) - Decision Tree (DT) - Random Forest (RF) - J48 (an implementation of C4.5) - Instance Based-k (IBk)	- UCSD-FICO dataset	- Organizing the original dataset and removing uncommon attributes - Reducing the original 20 attributes to 16 for analysis - Dataset Splitting	- Accuracy - Confusion Matrix - F1-Score - Kappa - MAE - MCC - Precision - Recall - RMSE	- Class Imbalance
(Alamri & Ykhlef, 2024)	To address the challenge of credit card fraud detection, focusing on the issues posed by highly imbalanced datasets.	Technical Paper	Credit card fraud detection	The proposed hybrid sampling method, which combines Tomek links for undersampling and BIRCH clustering with Borderline-SMOTE (BCBSMOTE) for oversampling, outperforms existing sampling techniques.	X			- Random Forest	- Synthetic dataset (PaySim)	- Data Cleaning - Exploratory Data Analysis (EDA) - Tomek Links to reduce undersampling - BCBSMOTE to enhance the representation of the minority class	- Accuracy - AUPRCC - Confusion Matrix - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance - Missing values - Noisy data - Overfitting
(Jadhav et al., 2022)	To develop an advanced machine learning model using XGBoost and SMOTE analysis to improve the detection and reduction of fraudulent credit card transactions.	Technical Paper	Credit card fraud detection	The application of SMOTE significantly improved the model's ability to detect fraudulent transactions with fewer false negatives, which are more critical in fraud detection scenarios.	X			- XGBoost	- Fraud detection dataset of European cardholders from Kaggle	- Scaling - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem	- Accuracy - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance
(Kumar et al., 2023)	To develop and evaluate a comprehensive fraud detection and prevention system using advanced machine learning techniques, with a focus on various sectors such as financial, e-commerce, and healthcare.	Technical Paper	Transaction fraud detection	The proposed fraud detection system, leveraging machine learning models, particularly logistic regression and XGBoost, along with SMOTE for handling class imbalance, achieves high accuracy in identifying fraudulent transactions.	X			- Logistic Regression (LR) - Random Forest (RF) - Support Vector Machine (SVM) - XGBoost	- Diverse datasets across different sectors	- Data Cleaning - Normalization of numerical features - SMOTE (Synthetic Minority Over-sampling Technique) to address the imbalance problem	- Accuracy - False Positive Rate (FPR) - True Negative Rate (TNR) - ROC	- Class Imbalance - Real-Time Processing
(Fatima et al., 2021)	To address the challenges of class-imbalanced learning, particularly focusing on minimizing the overlap degree in datasets to improve classification performance.	Technical Paper	Credit card fraud detection	The proposed algorithms, RONS (Reduce Overlapping with No-sampling), ROS (Reduce Overlapping with SMOTE), and ROA (Reduce Overlapping with ADASYN), effectively reduce the overlap degree in class-imbalanced datasets, thereby improving the performance of classification models.	X			- Support Vector Machine (SVM) - Logistic Regression (LR)	- Four UCI Machine Learning Repository datasets	- Cleaning and Normalization to remove any inconsistencies and normalized to ensure uniformity - Sparse Feature Selection - RONS (Reduce Overlapping with No-sampling) - ROS (Reduce Overlapping with SMOTE) - ROA (Reduce Overlapping with ADASYN)	- F1-Score - Gmean - Precision - Recall	- Class Imbalance - Class Overlapping
(Almazroi & Ayub, 2023)	To develop and evaluate the ResNeXt-embedded Gated Recurrent Unit (GRU) model (RXT) optimized with the Jaya algorithm for real-time	Technical Paper	Transaction fraud detection	The RXT-J model significantly outperforms traditional and contemporary algorithms, offering robust fraud detection capabilities with enhanced	X			- RXT-J (ResNeXt architecture and Gated Recurrent Unit (GRU), fine-tuned using the Jaya	- PaySim Synthetic dataset - Fraud detection dataset of European cardholders from	- Advanced Ensemble Feature Extraction (EARN) for advanced ensemble feature extraction - SMOTE (Synthetic	- Accuracy - AUC - F1-Score - Log Loss - MCC	- Class Imbalance - High Feature Dimensionality - Overfitting

	financial transaction fraud detection, aiming to enhance accuracy and efficiency in identifying fraudulent activities.			computational efficiency, making it highly effective for real-time applications.				optimization algorithm)	Kaggle - UCI Credit Card dataset	Minority Over-sampling Technique) to address the imbalance problem	- Precision - Recall - ROC	- Real-Time Processing
(Vorobyev & Krivitskaya, 2022)	To improve the efficiency of bank fraud detection systems by reducing the false positive rate (FPR) through an automatic rules generation framework that combines statistical and expert-based approaches.	Technical Paper	Transaction fraud detection	The proposed framework combining statistical and expert-based approaches effectively reduces false positives in fraud detection. The rules generated using tree-based models are interpretable and can be integrated into existing fraud detection systems. The approach proved to be satisfactory in a real industrial setting.	X			- Decision Tree - Random Forest - Gradient Boosting	- Dataset from a large bank	- Systematic sampling of transactions - Filtering noisy data - Removal of noisy labels, mistakes, and repeated transactions - Features are derived from the conditions and expressions of existing expert rules	- Accuracy - AUC PR - False Positive Rate (FPR) - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance - Dynamic Fraudulent Behavior - High False Positives - Interpretability of Models - Noisy data - Scalability
(Xie et al., 2023)	To develop a novel model that extracts transactional behaviors and learns behavioral representations from users' historical transactional behaviors for credit card fraud detection and behavioral changes caused by different time intervals between consecutive transactions.	Technical Paper	Credit card fraud detection	The proposed TAI-LSTM model outperforms conventional models in detecting fraudulent transactions, showcasing a better understanding of user behavior through attention mechanisms and interaction modules.	X			- TAI-LSTM (Time-Aware Attention-Based Interactive LSTM)	- Dataset from a Chinese financial institution - Fraud detection dataset of European cardholders from Kaggle	- Data Cleaning - Transaction Grouping - Timestamp Conversion to calculate time intervals between consecutive transactions - Transactional Embedding that includes the transaction itself and its historical context	- AUC - F1-Score - Precision - Recall	- Dynamic Fraudulent Behavior - Dynamic User Behavior - High False Positives - Ineffective Utilization of Temporal Patterns
(Chougule et al., 2024)	To explore the application of Amazon Web Services (AWS) SageMaker, a comprehensive machine learning platform, in developing and deploying an effective fraud detection system for credit card transactions.	Technical Paper	Credit card fraud detection	The deployment of ML models using AWS SageMaker for credit card fraud detection is highly effective, with high precision and recall. The platform's scalability and cost-efficiency make it ideal for real-time, large-volume transaction monitoring.	X			- Random Forest (RF) - Gradient Boosting - Deep Learning models	- Real-world dataset of credit card transactions	- Data cleaning - Normalization - One-Hot encoding - Temporal Features Extraction - Resampling	- Accuracy - F1-Score - Precision - Recall - ROC AUC	- Class Imbalance - Ineffective Utilization of Temporal Patterns - Real-Time Processing - Scalability
(Iscan et al., 2023)	To detect fraudulent activity on e-wallet platforms using machine learning techniques, with a focus on reducing false alarms.	Technical Paper	E-wallet fraud detection	LightGBM significantly reduced false alarms from 13,024 to 6,249. Achieved high detection accuracy, supporting small fraud detection teams by optimizing worker utilization by 52%.	X			- LightGBM (Light Gradient Boosting Machine) - XGBoost - Random Forest (RF) - Support Vector Machine (SVM - Multi-layer Perceptron (MLP)	- Dataset from Turkey's leading e-wallet platform	- Variable Selection - Aggregation - Feature Creation - Data Cleaning - Normalization - Encoding	- False Positives (FP) - False Positive Rate (FPR) - ROC AUC - True Positives (TP) - True Positive Rate (TPR)	- Class Imbalance - High False Positives
(W. Zhang et al., 2023)	To develop an effective credit card fraud detection model that overcomes the limitations of traditional methods by using a Generative Adversarial Network (GAN). This model, referred to as	Technical Paper	Credit card fraud detection	The proposed CCFD-GAN model outperformed classical credit card fraud detection models on various datasets, proving the effectiveness of the employed techniques.		X		- CCFD-GAN (Generative Adversarial Network)	- Fraud detection dataset of European cardholders from Kaggle - European credit card transaction dataset	- Neighbor Aggregation to generate new features - Sliding-Window Aggregation to capture temporal dependencies - Multi-Scope Outlier	- Accuracy - F1-Score - Precision - Recall	- Class Imbalance - Feature Space Limitation

	CCFD-GAN, aims to detect fraudulent transactions without needing fraud transaction samples for training.								- Online payment dataset	Degrees - User Grouping		
(Peng et al., 2023)	To develop a one-class adversarial fraud detection model to enhance the accuracy and stability of fraud detection using only normal user data.	Technical Paper	Credit card fraud detection	The proposed CS-OCAN model significantly improves the detection accuracy and stability of fraud detection compared to state-of-the-art one-class classification models.			X	- CS-OCAN (One-class adversarial fraud detection nets with class specific representations)	- UMD Wikipedia dataset - Credit Card Fraud Detection (CCF) dataset	- Autoencoder for Feature Extraction - Preliminary Feature Extraction - Pseudo Fraud Data Generation	- Accuracy - F1-Score - Precision - Recall	- Class Imbalance - Dynamic Fraudulent Behavior - Limited Data
(Xiang et al., 2023)	To develop a semi-supervised graph neural network for credit card fraud detection that can effectively utilize both labelled and unlabeled transaction data.	Technical Paper	Credit card fraud detection	The proposed Gated Temporal Attention Network (GTAN) significantly outperforms existing state-of-the-art methods in fraud detection tasks. This improvement is demonstrated across three different datasets, showing that GTAN is highly effective even with a small proportion of labelled data.			X	- Gated Temporal Attention Network (GTAN)	- Real-world transaction dataset (FFSD) - YelpChi dataset - Amazon dataset	- Attribute Embedding - Graph Construction to capture interactions among transactions	- AUC - Averaged Precision (AP) - F1-Score	- Complex Data Relationships - Limited Labels - Underutilization of Unlabeled Data

Ref	**Research Objective**	**Paper Type**	**Fraud Detection Focus**	**Main Conclusions**	**Methodology**	**Theoretical Framework**	**Problem(s) addressed**
(Nkomo & Breetzke, 2020)	To evaluate current credit card fraud detection methods used by banks, identify their difficulties, and propose a model incorporating artificial intelligence, geolocation, and data mining to improve fraud detection accuracy and efficiency.	Conceptual/ Review Paper	Credit card fraud detection	- The paper identifies current fraud detection methods such as Support Vector Machine (SVM), decision trees, self-organizing maps, artificial immune systems (AIS), and Bayesian networks, highlighting their respective strengths and weaknesses - A conceptual model using AI, data mining, and geolocation is proposed to improve the accuracy and efficiency of fraud detection. - The proposed model aims to reduce false positives and enhance the detection of fraudulent transactions by learning user behaviour patterns in real-time	- The study relies on secondary research, drawing information from books, research papers, journal articles, conference papers, organization websites, and academic websites - The research adopts an interpretivist approach to analyse and synthesize existing literature on credit card fraud detection methods	- Lawson's model for command and control to frame the decision-making process in fraud detection - Model is adapted to illustrate the flow of information from transaction initiation to the authorization and settlement phases, emphasizing the role of AI in improving decision accuracy	- The paper provides a thorough review of existing credit card fraud detection methods and their limitations - Introduction of a conceptual model incorporating AI, geolocation, and data mining for enhanced fraud detection - Emphasis on real-time learning of user spending patterns to improve the detection accuracy of fraudulent transactions - Highlighting the importance of feature selection in reducing noise and improving the efficiency of fraud detection models
(Cherif et al., 2023)	To provide a comprehensive review of the research on detecting and predicting fraudulent credit card transactions conducted from 2015 to 2021. The study aims to analyse the existing methods, highlight current research issues, and suggest future research directions in the context of emerging disruptive technologies.	Review Paper	Credit card fraud detection	- There is a limited investigation into the use of deep learning for credit card fraud detection, indicating a need for more research in this area - The study highlights the challenges associated with using new technologies like big data analytics and cloud computing in credit card fraud detection - The paper identifies several future research directions, including the need for more comprehensive studies on deep learning, addressing class imbalance, and exploring the integration of big data and real-time fraud detection systems	- The study follows a systematic review approach, analyzing 40 relevant articles published between 2015 and 2021 - The sources of information include books, research papers, journal articles, and conference papers - The selected papers were analyzed based on the publication date, coverage, and topics discussed, including the machine learning methods used, class imbalance issues, and feature engineering	- Not specifically mentioned	- The study provides a detailed survey of recent research on credit card fraud detection, focusing on the application of machine learning and new technologies - It highlights the main challenges faced in credit card fraud detection, including data-related, security-related, and implementation challenges - Providing a roadmap for future research directions to enhance credit card fraud detection systems.

(Priya & Saradha, 2021)	To evaluate the effectiveness of machine learning algorithms in detecting and preventing digital fraud across various sectors. It aims to highlight the need for a centralized fraud management platform where organizations can share fraud patterns and collaborate to improve fraud detection mechanisms.	Review Paper	Transaction fraud detection	- Traditional, siloed approaches to fraud detection are insufficient to combat modern fraud tactics - Machine learning algorithms can significantly improve the detection and prevention of fraud by learning from historical data and detecting complex patterns - A centralized fraud management platform is essential for organizations to share fraud-related data and collectively improve fraud detection capabilities - Organizations must adopt a proactive, rather than reactive, approach to fraud prevention	- Reviewing existing literature on fraud detection and prevention using machine learning - Analyzing the strengths and weaknesses of current fraud detection systems - Proposing a new model that incorporates machine learning algorithms for real-time fraud detection and prevention - Emphasizing the need for a centralized database where organizations can share fraud patterns and detection techniques	- The theoretical framework is based on the application of machine learning and artificial intelligence in fraud detection	- Highlighting the limitations of traditional fraud detection methods - Demonstrating the potential of machine learning algorithms in enhancing fraud detection accuracy and efficiency - Proposing a centralized fraud management platform for global collaboration - Suggesting a comprehensive life cycle approach to fraud detection and prevention
(Turksen et al., 2024)	To explore the legal implications and challenges of integrating AI and ML systems for automated suspicious transaction monitoring in the banking sector, with a focus on enhancing the integrity of AI systems. The study aims to identify the barriers to trust in AI among prudential banking supervisors and to investigate the drivers for adopting explainability technologies to increase transparency and understanding of complex algorithms.	Qualitative/ Research paper	Transaction monitoring	- AI/ML systems are increasingly important for banks, but their adoption is hindered by cost and economic efficiency considerations - There is a significant need for transparency and explainability in AI systems to gain trust from regulatory authorities - The complex and evolving regulatory landscape, particularly in the context of AML/CFT, poses significant challenges for the implementation of AI/ML systems - Banks are willing to adopt AI/ML technologies but face practical and regulatory barriers that need to be addressed	- The study uses a qualitative research design, including doctrinal legal research and semi-structured interviews - Interviews were conducted with senior managers from banks and IT companies in Ukraine, Estonia, and Poland - Data was collected through 12 interviews and analyzed to identify key themes and challenges related to the use of AI/ML in banking	- Based on the principles of AI ethics, transparency, explainability, and the regulatory requirements for AML/CFT	- Identifying the key challenges and barriers to the adoption of AI/ML systems in banking - Highlighting the importance of explainability and transparency in AI systems for regulatory compliance - Providing insights into the perspectives of bank managers and IT developers on the use of AI/ML for AML/CFT - Offering recommendations for improving AI/ML systems to enhance compliance and mitigate risks
(Ahmadi, 2024)	To investigate the impact of OpenAI and advanced machine learning technologies on enhancing fraud detection and prevention in the financial industry.	Review Paper	Transaction fraud detection	- OpenAI and machine learning technologies can significantly improve fraud detection by identifying complex and evolving fraudulent patterns - AI-based fraud detection systems can reduce false positives, enhance customer satisfaction, and protect financial institutions from financial losses - Generative AI, despite its potential for detecting fraud, also poses risks as it can be misused by fraudsters to perpetrate scams	- Review of existing literature and case studies - Analysis of current trends and technological advancements in fraud detection using AI - Examination of practical applications and real-world examples, such as the use of OpenAI by Stripe and Mastercard	- Grounded in the concepts of machine learning, neural networks, and generative AI, focusing on their application in fraud detection and prevention	- Demonstrating the potential of generative AI to enhance fraud detection capabilities by creating synthetic data for training models - Discussing various machine learning algorithms (e.g., logistic regression, decision trees, random forests, neural networks) used in fraud detection - Highlighting the successful implementation of AI by companies like Stripe and Mastercard to improve fraud detection and customer protection
(Sasmal, 2021)	To explore the use of Artificial Intelligence (AI) in preventing and mitigating card fraud and scams, and to assess the effectiveness of AI-based approaches compared to traditional methods.	Research paper	Card fraud detection	- AI provides a robust solution for detecting and mitigating card fraud through advanced data analysis, pattern recognition, and real-time decision-making - Traditional fraud detection methods are inadequate against sophisticated and evolving fraud tactics - AI-based approaches, while effective, come with challenges such as the need for extensive datasets, potential biases, and privacy concerns	- Analysis of current trends and the impact of AI in fraud detection - Examination of various AI models and techniques (supervised and unsupervised learning, anomaly detection, and behavioral analytics) - Consideration of ethical issues and challenges related to AI implementation	- Grounded in machine learning and AI principles, focusing on their application in financial fraud detection and prevention	- Highlighting the limitations of traditional rule-based fraud detection systems - Demonstrating the effectiveness of AI in identifying complex and evolving fraud patterns - Addressing the ethical considerations in using AI for fraud prevention - Proposing future directions for research, including the integration of blockchain and federated learning

				- Future research should focus on integrating blockchain, federated learning, and adaptive AI strategies to enhance fraud prevention			
(Dayyabu et al., 2023)	To investigate the application of artificial intelligence (AI) techniques in effectively and efficiently detecting credit card fraud and identifying fraudulent financial transactions.	Quantitative Study	Credit card fraud detection	- AI techniques, particularly machine learning and data mining, have a significant positive impact on credit card fraud detection - Fuzzy logic, although beneficial, is less utilized due to its lower accuracy compared to machine learning and data mining - AI provides better accuracy and efficiency in detecting fraudulent transactions compared to traditional methods	- Data collected from 100 respondents in the accounting and finance industry - Analysis using SPSS, including regression analysis, Pearson correlation coefficient, and reliability analysis	- Grounded in the application of AI techniques, focusing on machine learning, data mining, and fuzzy logic in the context of financial fraud detection	- Empirical evidence supporting the positive relationship between AI techniques and fraud detection accuracy - Highlighting the need for adopting advanced AI techniques over traditional method - Addressing limitations and suggesting improvements for future research
(Cirqueira et al., 2021)	To develop and evaluate design principles that align fraud experts' tasks with explanation methods (EM) for Explainable AI (XAI) decision support in fraud detection.	Empirical Paper	Fraud detection	- The developed design principles significantly enhance the alignment between fraud experts' tasks and XAI explanations, increasing trust and efficiency in fraud detection - Experts found the principles valuable, highlighting their utility in understanding AI predictions and aiding in decision support for fraud detection - The principles provide a structured approach to embedding user-centric XAI in fraud detection systems, thus fostering trust and improving the decision-making process	-Design science research methodology, involving iterative development and evaluation of design principles with feedback from fraud experts - Evaluation using an information quality framework through expert interviews and a simulation on a real transaction fraud dataset	- Combining Information Systems (IS) and Human-Computer Interaction (HCI) theoretical lenses to develop user-centric design principles for XAI in fraud detection	- Developing a set of design principles for aligning fraud experts' tasks with XAI explanations - Providing empirical evidence of the principles' utility through qualitative feedback from fraud experts and quantitative simulation results - Highlighting the importance of user-centric XAI in fostering trust and improving decision support systems for fraud detection
(Wolfsberg Group, 2022)	To provide a set of principles (the Wolfsberg Principles) that guide financial institutions (FIs) in the responsible use of Artificial Intelligence (AI) and Machine Learning (ML) in financial crime compliance, ensuring ethical considerations and effective risk management.	Position Paper	Financial crime risks	- Financial institutions should adopt a risk-based approach to AI/ML implementation, considering data ethics and ensuring fair, effective, and explainable outcomes - The responsible use of AI/ML can enhance the detection and management of financial crime, but requires careful oversight, transparency, and ethical considerations	- Conceptual framework development based on extensive regulatory, industry, and academic resources on data ethics - The development of guiding principles for FIs to manage the operational and reputational risks associated with AI/ML in financial crime compliance	- The principles are underpinned by existing literature and regulatory guidelines on data ethics and AI/ML applications in financial services - Integration of ethical and operational risk assessments into AI/ML governance frameworks within Fis	- Introduction of the Wolfsberg Principles, which provide a comprehensive guide for FIs on the ethical and effective use of AI/ML in financial crime compliance - Emphasis on the importance of transparency, accountability, and technical expertise in the deployment of AI/ML technologies - Advocacy for a balanced approach that mitigates risks while leveraging the benefits of AI/ML for fraud detection and prevention

AUC = Area Under the Curve
AUC PR = Area Under the Precision-Recall Curve
AUPRCC = Average precision and recall curve
MCC = Matthews Correlation Coefficient
PRC = Precision-Recall Curve
ROC = Receiver Operating Characteristic
ROC AUC = Receiver Operating Characteristic Curve Area Under Curve
Used for requirements