Journel/ Conference	Year of Publishing	Objective	Dataset Link	Algorithm/Technology Used	Advantages	Limitations
Credit Card Fraud Detection Using Machine Learning	2020	Detect credit card fraud using machine learning algorithms and compare the performance of Random Forest and Adaboost	Kaggle credit card fraud dataset, 2013	Random Forest, Adaboost	Random Forest reduces overfitting, high accuracy; Adaboost boosts weak classifiers	Adaboost is sensitive to noisy data and outliers
Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms	2022	To improve the accuracy of credit card fraud detection using state-of-the-art deep learning algorithms.	European card benchmark dataset	Convolutional Neural Networks (CNN), Decision Tree, Random Forest, SVM, Logistic Regression, XGBoos	Achieved high accuracy of 99.9%, with F1-score of 85.71%, precision of 93%, and AUC of 98%. Also addressed class imbalance using data balancing techniques.	Class imbalance remains an ongoing challenge. The model's performance decreases when applied to unseen data.
Credit Card Fraud Detection Using Machine Learning: A Comparative Study of Ensemble Learning Algorithms	2023	To identify fraudulent credit card transactions using machine learning algorithms and compare their performances	Kaggle Credit Card Fraud Dataset	Decision Tree, Logistic Regression, Support Vector Machine, Naive Bayes, XGBoost, Random Forest, Voting, Gradient Boosting, AdaBoost, Stacking	High accuracy with Random Forest and XGBoost; real-time detection of fraud	Naive Bayes has a high false positive rate, leading to inconvenience and extra workload for fraud analysts
Credit Card Fraud Detection	2024	To detect credit card fraud by using machine	A dataset from Kaggle consisting of	Random Forest, Logistic Regression,	Increased fraud detection	Limitations of individual algorithms,

		learning algorithms and assess their performance.	284,808 samples, 492 of which are fraudulent. The dataset contains 28 features, including the target variable..	Naive Bayes, Support Vector Machine (SVM), and Ensembling techniques to combine strengths of multiple models	accuracy (up to 96.60% for Naive Bayes + Logistic Regression)	risk of overfitting due to oversamplin g with SMOTE, and imbalanced datasets
Credit Card Fraud Detection using Deep and Machine Learning	2022	To compare different machine learning models and identify the best-suited model for detecting fraudulent credit card transactions	Kaggle credit card fraud dataset with 284,807 transactions (0.172% fraudulent)	Logistic Regression, XGBoost, Multi-Layer Perceptron (MLP)	MLP outperforme d other models, achieving higher accuracy, F1 score, recall, and AUC compared to XGBoost and Logistic Regression.	Dataset used is heavily imbalanced; model performance might not generalize well without addressing this issue extensively.
Supervised Machine Learning Algorithms for Credit Card Fraud Detection: A Comparison	2020	Evaluate an imbalanced dataset with the help of various supervised machine learning models to determine the best-suited algorithm for credit card fraud detection.	Publicly available dataset of European cardholders containing 284,807 transactions, out of which 492 are fraudulent (imbalanced dataset).	Decision Tree, k-Nearest Neighbor (kNN), Logistic Regression, Random Forest, Naive Bayes	1. Provides insights on the performance of various supervised learning models. 2. Decision Tree chosen for minimal prediction time. 3. Random Forest provides high accuracy. 4. kNN achieves the highest sensitivity. 5. Logistic Regression and Naive	1. Dataset imbalance affects model performance . 2. High computation al time for certain models (e.g., kNN). 3. Naive Bayes has low sensitivity compared to others. 4. Logistic Regression is limited in handling non-linear data. 5. Time

					Bayes show strong performance in precision under certain threshold settings.	taken for prediction varies significantly across models, impacting real-time detection suitability.
Research on Credit Card Fraud Detection Model Based on Distance Sum	2009	The objective is to detect fraudulent transactions effectively, improving over traditional methods.	It is described as real credit card data from a domestic commercial bank, consisting of 16,584 transaction records, of which 1,449 are fraudulent and 15,135 are non-fraudulent	Outlier mining algorithm based on distance sum, using Euclidean distance to detect fraud by identifying outliers	Effective at detecting fraud, especially when fraudulent transactions are far fewer than normal transactions.	The method's performance depends on the choice of threshold for outlier detection and data standardizati on
Hybrid Multi-Level Credit Card Fraud Detection System by Bagging Multiple Boosted Trees (BMBT)	2017	To propose an ensemble model, Bagging Multiple Boosted Trees (BMBT), to effectively detect fraudulent transactions in imbalanced credit card data by overcoming data hugeness and imbalance	The dataset used is from the *UCSD – FICO Datamining Contest 2009* (Hard Version T2)	The BMBT model uses an ensemble method combining bagging with boosted decision trees. It is implemented on the Spark architecture for big data stream processing	Effective in dealing with imbalanced data. High performance levels in terms of AUC, BCR, and BER when compared to existing models.	The model still shows moderate false positive rates (FPR levels of 0.23), which could be reduced in future work
Performance Evaluation of Machine Learning Algorithms for Credit Card Fraud Detection	2019	The paper evaluates the performance of various supervised and unsupervised machine learning	Credit Card Fraud Detection Dataset	Unsupervised Learning Algorithms: Self-Organizing Maps, K-means, Isolation Forest, Local Outlier Facto	Random Forest: Robust to noise and outliers. Neural Networks (NN):	Showed poor results in detecting true positives. And Certain algorithms like SVM

		algorithms in detecting credit card fraud in highly imbalanced datasets		Supervised Learning Algorithms: Random Forest, Neural Networks, Deep Learning	Effective in recognizing patterns. Deep Learning (DL): Efficient in high-dimensional spaces.	and NN showed NaN values where they failed to detect positive or negative instances correctly