Credit Card Fraud Detection Using Machine Learning Techniques

University of San Diego

Course: AAI-590 — Capstone Project

Group 02

Teammates: Angshuman Roy, Harish Kapettu Acharya, Sandeep Kumar Jakkaraju

Instructor: Zahid Wani

Date: November 2025

Introduction

Credit card fraud is a high-impact problem in financial services, resulting in direct financial losses, operational risk, and reputational challenges. Detecting fraud is difficult because fraudulent behavior is rare, adaptive, and often hidden within large volumes of legitimate transactions. The Kaggle Credit Card Fraud Detection dataset includes 284,807 transactions with only 492 fraud cases (0.172%), creating an extremely imbalanced binary classification task. This draft describes the experimental methods,model architectures, and comparative modeling plan to evaluate both traditional machine-learning models and deep learning architectures

Findings and Model Evaluation

Model Learning Effectiveness and Evidence of Overfitting/Underfitting

Across all supervised learning algorithms implemented—Logistic Regression, Random Forest, SVM, Naïve Bayes, k-NN, XGBoost, LightGBM, CatBoost, and a Deep Neural Network (DNN)—the models demonstrated strong capability to learn the fraud-detection task despite the extreme class imbalance. Training and validation performance indicate no major signs of overfitting for tree-based models or the DNN. Regularization mechanisms inherent in boosting algorithms prevented overfitting, while the DNN showed steadily decreasing loss with early stopping, indicating stable optimization.

Models such as Logistic Regression, Naïve Bayes, and Linear SVM tended to underfit, as evidenced by low precision despite high recall. Tree-based boosting models—especially XGBoost and LightGBM—demonstrated the most effective learning performance with strong F1-scores and AUC metrics. Unsupervised anomaly models such as Isolation Forest, LOF, and Autoencoders underperformed relative to supervised methods.

Support for the Research Question and Hypothesis

The research question asked how effectively machine learning methods can detect fraudulent transactions despite severe dataset imbalance. Empirical results strongly support this hypothesis. XGBoost, LightGBM, and the DNN achieved recall scores above 0.77, precision values above 0.89, and ROC–AUC values near 0.98. These results demonstrate strong discriminatory power between fraudulent and legitimate transactions.

Problem-Solving Impact and Application Relevance

Fraud detection systems require high recall to prevent missed fraud cases and high precision to reduce customer friction and analyst overload. The strongest models—XGBoost, LightGBM, and the DNN—achieved an optimal balance between these goals, making them suitable for operational deployment. Their computational efficiency and scoring stability align with industry-standard fraud detection requirements.

Unexpected Results and Observations

Several unexpected findings emerged during experimentation. Linear models exhibited high recall but extremely low precision, making them impractical despite detecting many fraud cases. CatBoost performed worse than expected given its reputation in tabular data problems, likely due to the dataset's PCA-transformed numeric features. Unsupervised anomaly detection approaches performed poorly, reinforcing the advantage of supervised learning when labeled data is available.

Next Steps for Future Work

Future work includes advanced hyperparameter optimization, enhanced feature engineering, and the development of ensemble model stacking strategies. Incorporating graph-based models such as Graph Neural Networks may capture relational fraud patterns. For productionization, the model would require a real-time scoring API, drift monitoring, continuous retraining pipelines, and integration with human-in-the-loop review systems.

Conclusion

The results demonstrate that machine learning models—particularly XGBoost, LightGBM, and DNNs—are highly effective at detecting fraudulent transactions under severe class imbalance. The findings support the project's research hypothesis and identify clear paths for operational deployment and future research opportunities.