

Executive Summary: Credit Card Fraud Detection

The primary objective of the fraud detection model is to differentiate fraudulent transactions from legitimate ones using machine learning. The analysis consisted of data preprocessing and exploration, model development and evaluation and tableau dashboard creation. It highlights key performance metrics across multiple models and identifies influential features in detecting fraud.

The model evaluation using classification metrics (e.g., accuracy, precision, recall, and ROC-AUC) suggests that tree-based models like XGBoost, Random Forest, and CatBoost perform best. High recall is critical in fraud detection, ensuring fraudulent transactions are captured. However, a trade-off exists with precision, meaning some false positives occur. Logistic Regression and LDA show lower overall effectiveness, particularly in capturing fraud cases. Catboost model provide to be the best one overall with 100% accuracy, 83% f1-score, 93% precision, 76% recall and 98% AUC-ROC.

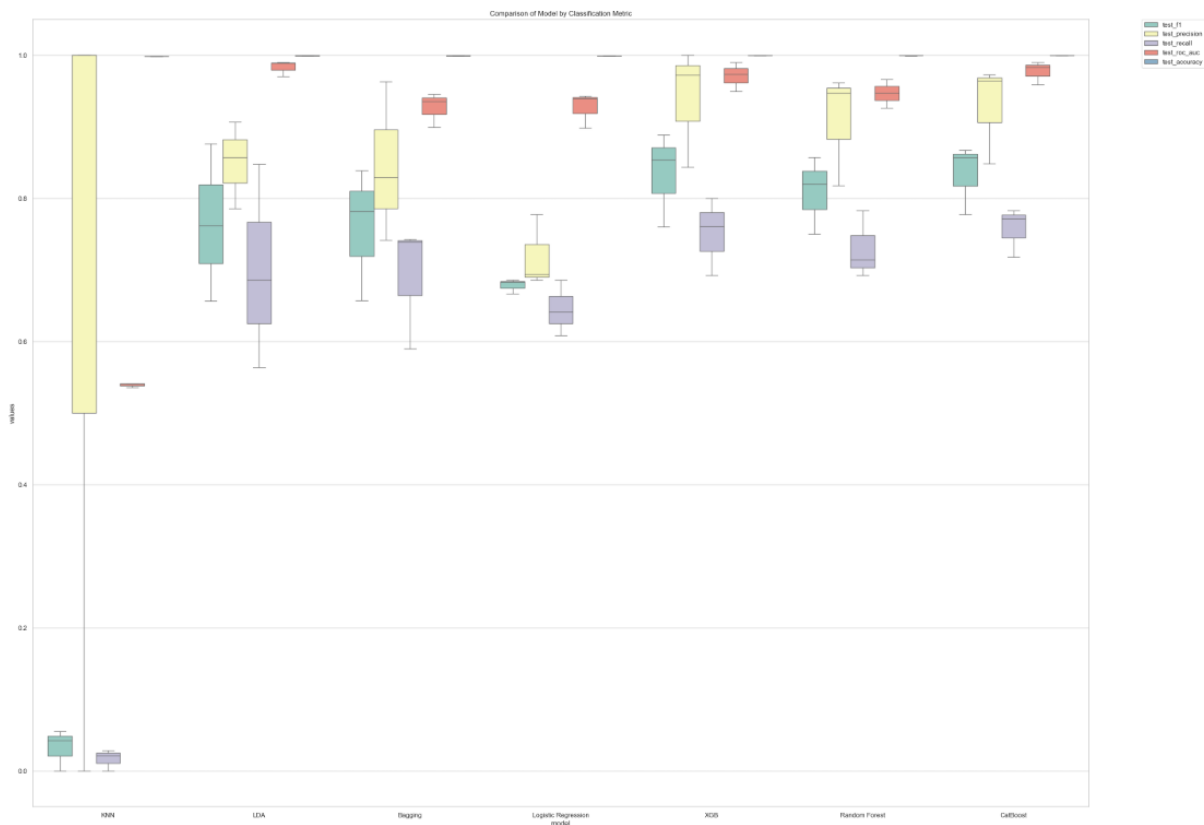


Image 1: Model comparison

SHAP value analysis indicates the following top contributors to fraudulent activity. V14, V12, V1 and V8 features exhibit strong predictive power, with lower values correlating with fraudulent transactions. V4, V11 and V21 have strong predictive power, with higher values correlating with fraudulent transactions. Amount of transaction also has a significant impact on the likelihood of fraud according to SHAP values.

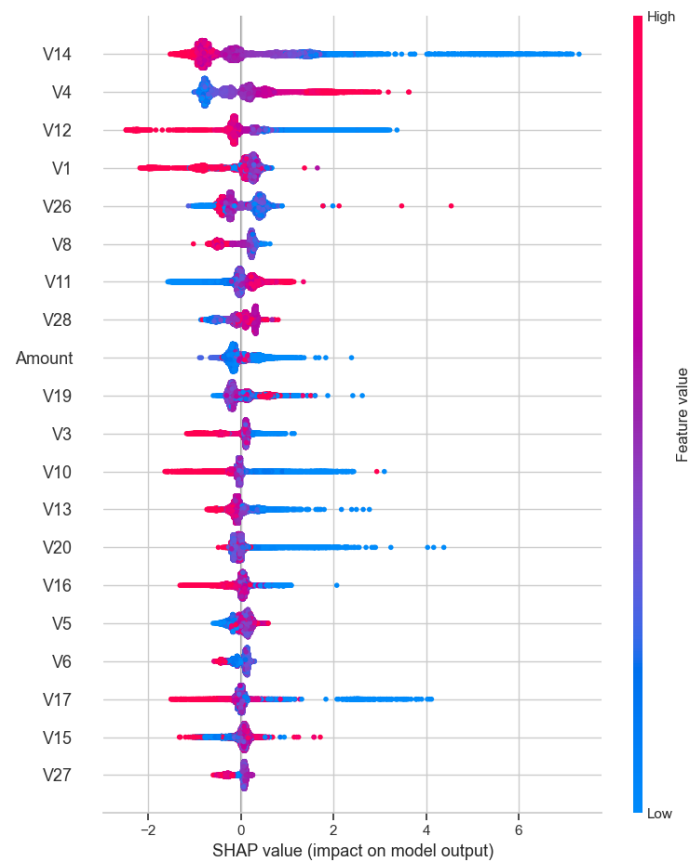


Image 2: SHAP feature importance

Recommended actions would be to enhance real-time monitoring. Put our implemented real-time fraud detection CatBoost model with optimized parameters into production. Utilize high-impact features mentioned to differentiate between fraudulent and legitimate transitions. Regularly retrain the model on new transaction data to adapt to emerging fraud patterns.

By leveraging the model's insights, fraud detection can be both accurate and proactive, reducing financial losses while minimizing legitimate transaction disruptions.