VIT-AP
UNIVERSITY
Apply Knowledge. Improve Life!®

Vellore Institute of Technology Amaravathi - 522237
Batch - 2022 (SEMESTER – FALL SEM 2024 - 2025)

Theory Research Paper

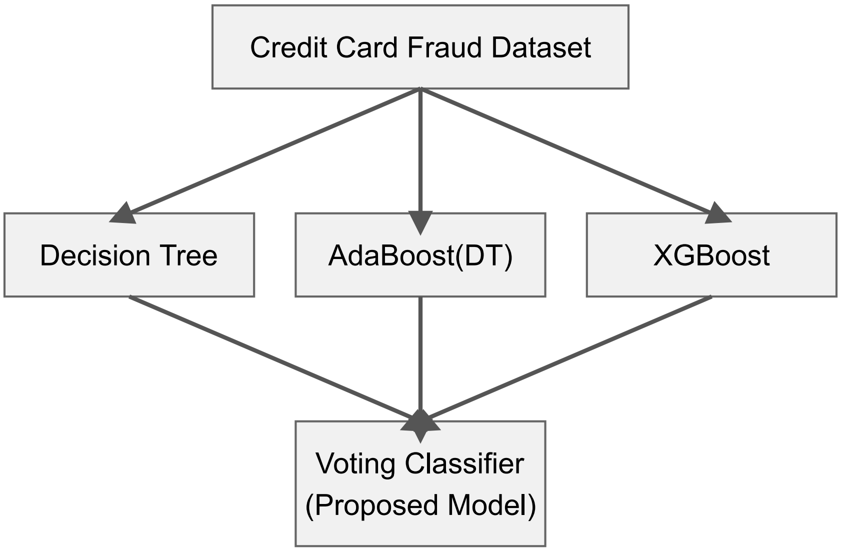
Prof. Srinivas Arukonda
Department of Computer Science and Engineering

Robust Approach for Detecting Fraudulent Transactions

Team :

NALLAGULA KARTHIK SAGAR	-	22BCE8811
JOEL BINU PHILIP	-	22BCE9316
TRILOK SHANKAR	-	22BCE20233
UDAY MADHAV	-	22BCE8458

Proposed Model Diagram



Contribution

In this study, we propose a novel ensemble model for the detection of fraudulent credit card transactions. Unlike traditional approaches that rely on a single machine learning algorithm, our proposed model utilizes a Voting Classifier that combines the predictions of multiple base models.

The base models used in our approach are:

1. **Decision Tree Classifier**: Decision trees are intuitive and easy to interpret, making them a popular choice for fraud detection tasks. They can capture complex relationships in the data and provide good performance on tabular datasets like credit card transactions.
2. **K-Nearest Neighbors (KNN)**: KNN is a non-parametric, instance-based learning algorithm that can effectively identify anomalies and outliers in the data. It is well-suited for detecting fraudulent transactions, which often exhibit patterns distinct from legitimate ones.

3. **AdaBoost Classifier**: AdaBoost is an ensemble technique that combines multiple weak learners to create a strong predictive model. By iteratively adjusting the weights of misclassified instances, AdaBoost can improve the model's ability to detect subtle fraudulent patterns.
4. **XGBoost Classifier**: XGBoost is a highly optimized and scalable implementation of gradient boosting, a powerful ensemble method. Its ability to handle complex, non-linear relationships in the data makes it a robust choice for fraud detection.

The key contribution of our work is the integration of these diverse base models into a Voting Classifier. By leveraging the strengths of each individual model, the Voting Classifier is able to make more accurate and reliable predictions compared to using a single model alone.

The Voting Classifier works by taking the majority vote or the average prediction of the base models to determine the final classification. This ensemble approach helps to mitigate the weaknesses of individual models and improves the overall robustness and generalization of the fraud detection system.

Furthermore, the use of a Voting Classifier aligns well with the complex and evolving nature of credit card fraud. As new fraud patterns emerge, the ensemble model can adapt and learn from the diverse perspectives of the base models, leading to better performance in detecting both known and novel fraudulent transactions.

By combining multiple state-of-the-art machine learning algorithms in a Voting Classifier, our proposed model offers a comprehensive and robust approach to credit card fraud detection. This contribution is expected to have a significant impact on the financial industry, helping to reduce the financial and reputational damages caused by fraudulent activities.

Results Comparison Table

Model	Accuracy	F1 Score
XGBoost	0.999391	0.834586
Decision Tree	0.999228	0.76595
KNN	0.998303	0.039735
AdaBoost	0.999181	0.753521
PROPOSED MODEL	0.999597	0.840149