

Datasheet and Model Card

Developed by: Naomi Jobson Mitchual

Data Source: The data used in this project is credit card transaction data.

Data Description: The dataset contains credit card transactions labeled as fraudulent (1) or not fraudulent (0). Each transaction is represented by a set of features derived from PCA (Principal Component Analysis) decomposition of the original data.

Data Preprocessing: The data was preprocessed by performing PCA to reduce dimensionality and retain the most significant components. The data was then split into training and testing sets using stratified means to ensure balanced representation of both classes in each set.

Model Card

Model Details: The models were developed by Naomi Jobson Mitchual to help in credit card fraud detection for a project undertaken for Professional Certification in Machine Learning.

Model Usage: The models are designed to assist in identifying fraudulent credit card transactions, thereby reducing the risk of fraud.

Models Trained:

- Logistic Regression Classifier
- Decision Trees
- Support Vector Machine (SVM)
- Random Forest

Data Used: The models were trained using credit card transactional data.

Metrics for Scoring: The performance of the models was evaluated using the following metrics:

- Accuracy
- F1 Score
- Precision
- Area Under the Curve (AUC)
- Recall

Ethical Considerations: Data was anonymized, and all Personally Identifiable Information (PII) was removed before model training to ensure privacy and ethical handling of the data.