CREDIT CARD FRAUD DETECTION

SUBMITTED BY: AMIT SAINI



MOTIVATION

Cybersecurity is becoming increasingly important. When it comes to digital security, the most difficult task is detecting unusual activities. Credit limit in credit cards sometimes helps us make purchases even if we don't have the amount at that time. These features are misused by cyber attackers, We need a system that can abort the transaction if it finds fishy

DATA PROCESSING & UNDERSTANDING

- The exact variables are not disclosed due to security concerns, however, they have been modified versions of PCA. As a consequence, there are one time, 29 feature columns and one final class column to be found.
- The dataset is imbalanced towards a feature "legit transaction".
- Our dataset has no null values
- The mean amount of Fraudulent transactions is greater than the legit
- We removed duplicate transactions

TRAINING DATA & TEST DATA - SPLITTING DATA

- Since our dataset is significantly unbalanced, we first undersample the data from the majority class.
- We upsample the minority class using SMOTE and build a sample dataset containing similar distribution of normal transactions and Fraudulent Transactions
- We divide the data into two datasets training data and testing data

MODEL BUILDING - RANDOM FOREST

	precision	recall	f1-score	support
0 1	0.89 0.99	0.99 0.87	0.94 0.93	148 148
accuracy macro avg weighted avg	0.94 0.94	0.93 0.93	0.93 0.93 0.93	296 296 296

MODEL BUILDING - ADABOOST

	precision	recall	f1-score	support
0 1	0.88 0.98	0.98 0.87	0.93 0.92	148 148
accuracy macro avg weighted avg	0.93 0.93	0.93 0.93	0.93 0.93 0.93	296 296 296

MODEL BUILDING - XGBOOST

	precision	recall	f1-score	support
0 1	0.89 1.00	1.00	0.94 0.94	148 148
accuracy macro avg weighted avg	0.95 0.95	0.94 0.94	0.94 0.94 0.94	296 296 296

REFERENCES AND RESOURCES



https://www.youtub
e.com/watch?
v=NCgjcHLFNDg

SOURCE DATASET

https://www.kaggle.com/mlg-ulb/creditcardfraud

WEBSITES

https://iitg.ac.in/sa/c aciitg/sa22/course/

THANKYOU