ML Classification Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1. **Abhishek Mishra (Abhishekmishra@gmail.com):**

Exploratory Data Analysis (EDA)

Data frame description
Data frame shape
Understanding column "SEX".
Understanding column "EDUCATION".
Understanding column "MARRIGE".

Feature Engineering

Used Histogram plot, Pair plot to understand the data.
Done some feature engineering on the data set.

Data Pre-processing

Use SMOTE Because the data is highly imbalanced.
.

Model Implementation

Logistic Regression with hyper parameter tuning.
Random Forest classification using hyperperameter tuning.
Random Forest Classification without hyperperameter tuning.
Decision Tree with hyper parameter tuning.

Creating Data Frame of all Evaluation Matrix with respect of models

Model Explainability

Model Explainability using shap

Conclusion from Model Training

2. **Kurva Mallesh (kurvamallesh36@gmail.com):**

Exploratory Data Analysis (EDA)

Data frame description

Data frame shape

Check the change status of the payment in months

Relation between credit limit and the default payment next month

some bill statement amounts greater than credit limit

Feature incoding And creation

Used Histogram plot, Pair plot to understand the data.

Done some feature engineering on the data set.

Model Implementation

XGBoost with hyper parameter tuning.

Logistic Regression with hyper parameter tuning.

Random Forest classification using hyperperameter tuning.

Creating Data Frame of all Evaluation Matrix with respect of models

Model Explainability

Model Explainability using shap

Conclusion from Model Training

4. **Arunesh Mishra(Arunesh12mishra@gmail.com):**

Exploratory Data Analysis (EDA)

Data frame description

Data frame shape

Relation between credit limit and the default payment next month

Customers who had no consumption in 6 months then

Feature Engineering

Used Histogram plot, Pair plot to understand the data.
Done some feature engineering on the data set.

Model Implementation

XGBoost with hyper parameter tuning

Logistic Regression with hyper parameter tuning.

Random Forest classification using hyperperameter tuning.

Creating Data Frame of all Evaluation Matrix with respect of models

Model Explainability

Model Explainability using shap

Conclusion from Model Training

GithubRepository link:

Abhishek Mishra- https://github.com/abhishekmishra-bareilly/ML-Classification-Capstone-project

Kurva Mallesh - https://github.com/kurvamallesh/Capstone-project-on-classification

Arunesh Mishra - https://github.com/kajuun/ML-classification-capstone-project

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

This project is aimed at predicting the case of customers default payments in Taiwan. From the perspective of risk management, the result of predictive accuracy of the estimated probability of default will be more valuable than the binary result of classification - credible or not credible clients.
This dataset contains information on default payments, demographic factors, credit data, history of payment, and bill statements of credit card clients in Taiwan from April 2005 to September 2005.
We are all aware what is credit card. It is type of payment payment card in which charges are made against a line of credit instead of the account holder's cash deposits. When someone uses a credit card to make a purchase, that person's account accrues a balance that must be paid off each month.

Credit card default happens when you have become severely delinquent on your credit card payments.Missing credit card payments once or twice does not count as a default. A payment default occurs when you fail to pay the Minimum Amount Due on the credit card for a few consecutive months.