Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Shubham Srivastava: **E-Mail-** shubham.mach30@gmail.com • Loading Dataset • Data Inspection • Exploratory Data Analysis • Evaluating various models • Conclusion
Please paste the GitHub Repo link.
Github Link:- https://github.com/Shubhamverse/Credit-Card-Default-Prediction-Supervised-ML-Classification **Drive Link:-** https://drive.google.com/drive/folders/1T7pb-wZxn5PG-1pzDqQDxf0IRCzbdXKw?usp=sharing
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

In this Supervised Classification Project, we were provided with dataset i.e. Default of Credit Card clients .
As the first step, perform Data Inspection after the loading of Dataset through which we get to know the summary of data, shape of data, null value count in data and detail about datatype of columns.

Next, we do exploratory analysis during which we apply label encoding on categorical columns such as sex, education, marriage. Then, apply SMOTE(Synthetic Minority Oversampling Technique) to remediate Imbalance in dependent column. After which we apply feature engineering on the data to acquire some meaningful features for better prediction accuracy. At last, we apply One Hot encoding on features such as sex(Male, Female).

Then, we apply various classification models where being defaulter or not is dependent variable while others are independent variable. We applied Logistic Regression, SVC, Decision Tree Classification, Random Forest Classification, XGBoost Classification get values for evaluation metrices.

Finally, we derive conclusion based on results shown through various Classification models evaluation metrices that Random Forest Classifier has highest Accuracy, ROC-AUC compared to other classifiers.