Title: Prediction of Credit Card Fraud
Subtitle: Capstone Project Report

1. **Data understanding and exploring**
2. **Exploratory Data Analysis (EDA)**
    - Handling Missing Values
    - Checking the distribution of the classes
    - Outliers treatment
    - Observe the distribution of classes with time
    - Observe the distribution of classes with amount
3. **Data Splitting**
    - Train/Test Split: Explaining how the data was split into training and testing sets.
    - Scaling the Data: Describing the normalization or standardization techniques applied.
4. **Data Preparation for Modeling**
    - Check and Mitigate Skewness: Explaining the techniques used to handle skewness in the data.
    - Handling Data Imbalance: Describing the methods used to balance the dataset (e.g., SMOTE, ADASYN). Handling data imbalance as we see only 0.17% records are the fraud transactions
5. **Model Building**
    - Algorithms Used: Train the model with various algorithm such as (Logistic Regression, Decision Tree, Random Forest, XGBoost etc.).
    - Hyperparameter Tuning: Describe the process of tuning hyperparameters using Grid Search Cross Validation and find the optimal values of the hyperparameters.
6. **Model Evaluation**
    - Evaluation Metrics : Explanation of why accuracy might not be suitable due to data imbalance and focus on Precision, Recall, and ROC-AUC. We have to look for a balance between Precision and Recall over Accuracy. We also have to find out the good ROC score with high TPR and low FPR in order to get the lower number of misclassifications.
    - Results: Present the evaluation results for the models, highlighting Precision, Recall, and ROC-AUC.
7. **Conclusion and Future Work**
    - Summary of Findings : Key takeaways include the effectiveness of balancing techniques and the performance of different models.

    - Future Work : Future steps could include trying more advanced balancing techniques and exploring deep learning models.