CAPSTONE PROJECT 2
Online Payments Fraud Detection Project Proposal
by Sushmitha Amirishetty PGA 36

1. Executive Summary:

Due to the rapid increase in online payments the online fraud has also drastically increased. It became very important for credit card companies to prevent this from happening, so our project aims to gain insight and to develop a model that provides predictions on fraudulent activities from online credit card transactions.

2. Problem Statement:

Background: Online payment systems have changed the way transactions are conducted, providing convenience and efficiency in payment process. However, with the widespread use of these systems, there has been a significant rise in payment fraud, particularly in transactions involving credit cards. As more consumers and businesses rely on online payments, the risks associated with fraudulent activities have also escalated. Detecting and preventing online payment fraud has thus become a critical priority for credit card companies to protect their customers from unauthorized charges and financial loss.

Objective: The objective of this project is to develop a model for detecting online payment fraud in credit card transactions. By analyzing transaction data and identifying patterns indicative of fraudulent behavior.

Scope: The project focuses on online payment fraud in credit card transactions. It involves analyzing, visualizing and processing online transactions, building a model that is able to detect fraud transactions with good accuracy and also to deploy the model.

3. Data Sources:

- We will use a dataset that contains information on name, payment type, transaction amount, balance, etc**.**

4. Methodology:

Data Integration:

- The data will be integrated into Jupiter notebook for preprocessing, building, evaluating and selecting the model.
- It will be integrated into gradio and hugging face for deployment.

Preprocessing:

- The dataset will be processed to handle missing values, duplicates, outliers and imbalances in the fraud and nonfraud classes. Visualization will also be done alongside with it. Feature selection, feature engineering, encoding and scaling will be performed when required.

Model Building:

- Classification models such as Logistic Regression, K Neighbors Classifier, SVC, Decision Tree, Random Forest, Extra Trees, AdaBoost, Gradient Boosting, XGB will be built along with ANN as the data set is big.

Model Evaluation and Selection:

- The models will be evaluated based on metrics such as accuracy, precision, recall, F1-score, confusion matrix, ROC curve and AUC.
- The models will be trained and tested on the dataset using cross-validation to ensure overall performance.
- The model that performs the best in the evaluation will be selected.

Model Deployment:

- After the model is evaluated then the model is deployed using gradio and hugging face.

5. Expected Outcomes:

- Expecting to successfully visualize the analysis and to gain meaningful insights out of it.
- Expecting to get a good prediction of the Fraud data.
- Expecting to successfully deploy the model.

6. Tools and Technologies:

- Jupyter notebook: For data preprocessing, exploratory data analysis.
- Gradio: Used for creating user interface
- Hugging Face: It is used for model deployment

7. Risks and Challenges:

- Data preprocessing: Making sure that the raw data is cleaned and preprocessed.
- Avoid overfitting: As the data is big, it’s a challenge to avoid overfitting.

8. Conclusion:

The proposed fraud detection project holds immense potential to suppress credit card fraud online. By leveraging data driven insights and predictive model the credit card companies can potentially take preventive measures to stop fraudulent activities and protect their customers in a better way.