ALY 6040:
DATA MINING APPLICATIONS

Assignment 6
Contribution to the Group Project: Online Payments Fraud Detection

Submitted To:
Dr. Chinthaka Pathum Dinesh, PhD, Prof. Herath Gedara,
Faculty Lecturer

Submitted By:
Abhilash Dikshit

Academic Term: Spring 2023
Graduate Student at Northeastern University, Vancouver, BC, Canada
Master of Professional Studies in Analytics

May 20, 2023

Title: Contribution to the Group Project: Online Payments Fraud Detection

Introduction:

This document outlines my significant contributions and roles as a coordinator, leader, and coding contributor to the successful completion of our group project on Online Payments Fraud Detection. Our project aimed to develop a robust fraud detection model utilizing the provided dataset on Kaggle. This report highlights my active involvement in coordinating the team, leading project activities, and contributing to the coding aspects.

Role within the Team (Coordinator, Leader, and Coding Contributor):

In our group, I assumed multiple roles, including coordinator, leader, and coding contributor. As a coordinator, I focused on facilitating effective communication and maintaining team cohesion. I ensured that all team members were informed about project progress, assigned tasks, and deadlines. Additionally, I coordinated regular meetings, provided guidance, and resolved any conflicts or issues that arose within the group.

Moreover, I took on the role of a leader, guiding the team through different stages of the project. I provided strategic direction, set project objectives, and ensured that we stayed on track towards meeting our goals. I also encouraged collaboration, fostered a positive team dynamic, and motivated members to give their best effort.

In terms of coding, I actively contributed to the project by writing code and implementing various components. I collaborated with team members to develop data preprocessing techniques, feature engineering methodologies, and model implementation. I leveraged my coding skills to integrate different functionalities, optimize code efficiency, and ensure seamless interactions between various project components.

Group Members and Their Roles:

Our group consisted of the following members:

1. Abhilash Dikshit: Data Preprocessing, Feature Engineering, Model Development and Evaluation, Report Writing
2. Minesh Patil: Exploratory Data Analysis (EDA) and Visualization
3. Milan Prajapati: EDA and Visualization, Model Development and Evaluation, Report Writing

4. Murtaza Vora: EDA and Visualization, Model Development and Evaluation, Report Writing
5. Shamim Sherafati: EDA and Visualization, Project Documentation and Report Writing

My Contribution and Analysis:

During the project, my responsibilities encompassed data preprocessing and feature engineering, including tasks such as data cleaning, handling missing values, and performing feature selection and transformation to ensure the dataset was suitable for analysis. Additionally, I conducted an analysis utilizing anomaly detection techniques. I implemented algorithms such as Isolation Forest and Local Outlier Factor (LOF) to detect potential outliers and anomalies within the dataset. The objective was to identify irregular patterns and potentially fraudulent transactions that may not be captured through conventional classification approaches. Furthermore, I focused on model development and evaluation, employing various machine learning algorithms such as logistic regression, random forest, and support vector machines (SVM) to train and evaluate fraud detection models. To assess the effectiveness of these models, I utilized performance metrics including accuracy, precision, recall, and F1-score.

Results and Incorporation into the Final Report:

The results of the analysis from over 6 billion rows and 11 attributes revealed promising findings, identifying transactions with suspicious behavior that were not detected by our primary classification models. However, due to time constraints and prioritization of other analysis approaches, we decided not to include the anomaly detection analysis in the final report. Random Forest and SVM gave a 100% accuracy based on our training and test data.

Conclusion:

In conclusion, my contributions as a coordinator, leader, and coding contributor significantly impacted the success of our group project on Online Payments Fraud Detection. Through effective coordination, leadership, and active participation in coding tasks, I helped drive the team towards achieving our objectives. Working collaboratively with other group members, we successfully completed the project, gaining valuable insights into fraud detection methodologies and fostering a strong team dynamic.