Credit Card Fraud Detection

Introduction to Data Science

Project Report

Course Code: **CSC-495**
Lab Instructor: **Engr. Ayesha Khan**
Course Instructor: **Dr. Hina Shakir**

Group Members

Muhammad Nadeem (Group Leader)	**02-131212-054**
Arsalan Jah	**02-131212-078**
Ali Haider	**02-131212-007**

Teacher Signature: _____
Remarks: _____
Submission Date: 9th January 2024

BAHRIA UNIVERSITY
KARACHI CAMPUS
Department of Software Engineering

ACKNOWLEDGEMENT

Besides the efforts of all team members, this thematic section of the project report depends largely on the encouragement and guidance of our teachers. We take this opportunity to express our greatest gratitude to **Engr. Ayesha Khan** and **Dr. Hina Shakir** who contributed to the approval of this subject of the project. We can't give them enough credit for their amazing support and help. They motivated and encouraged to devote a lot of time to the selection of the appropriate project topic. Without their encouragement and advice, we would not have been able to choose the right subject. The input and support received from all team members including Syed Muhammad Askari, Muhammad Zain, and Hamza Zafar is vitally important. The team spirit shown by each member allowed project relationship to work successfully.

Table of Contents

Introduction & Problem

The proposed project aims to develop a machine learning-based system for detecting credit card fraud using Python and popular machine learning libraries such as Scikit-learn. The system will be designed to identify fraudulent transactions in real-time by analyzing incoming credit card transactions.

2. Paradigms

The project will employ the following paradigms:

- Data preprocessing: Removing irrelevant or redundant information from the credit card transaction data.
- Feature selection: Identifying the most important features for detecting fraud.
- Model training: Training a machine learning model using the preprocessed data and selected features.
- Model evaluation: Evaluating the performance of the trained model using metrics such as accuracy, precision, recall, and F1 score.
- Real-time fraud detection: Detecting fraud in real-time by analyzing incoming credit card transactions.

Model & Explanation (CCN)

The system will use various machine learning algorithms for credit card fraud detection, including Decision Trees, Random Forest, Artificial Neural Networks, Naive Bayes, and Logistic Regression

. These algorithms will be trained and tested using large datasets to optimize their performance in detecting credit card fraud.

Libraries Used & Overview

The project leverages several libraries to achieve its objectives:

- **NumPy:** Utilized for numerical operations, providing efficient data handling and manipulation.
- **Pandas:** Used for data analysis and manipulation, ensuring effective handling of information.
- **SCIKIT:** used for various machine learning tasks, including classification, regression, and clustering
- **Matplotlib**:used for analysis purpose such as graphs
- **Pickle**:used for saving and loading the trained model

Conclusions

In conclusion, credit card fraud is a significant problem that can result in financial losses for both individuals and businesses. Machine learning and data science techniques, such as those discussed in the research paper "Credit Card Fraud Detection using Machine Learning and Data Science," can be used to develop effective fraud detection systems. By utilizing machine learning algorithms and libraries like Scikit-learn, developers can preprocess credit card transaction data, select important features, train and evaluate models, and implement real-time fraud detection systems. These systems can help prevent fraudulent transactions and protect individuals and businesses from financial losses.