

INTERNSHIP REPORT

Credit Card Fraud Detection



Submitted by:

Anuroop Arya

Offer Number : U241103483

Project Submission Number : UT240311447

Introduction:

Financial institutions face the persistent challenge of fraudulent credit card transactions, which pose a significant threat to financial assets and customer trust. To address this issue, the objective of this project is to implement a robust fraud detection system using machine learning techniques. The system aims to accurately identify and prevent fraudulent activities in real-time or near-real-time, thereby safeguarding the financial assets of the institution.

Objectives:

1. Develop a machine learning model for fraud detection: Utilize historical credit card transaction data to train a machine learning model capable of detecting fraudulent transactions.

2. Identify features indicative of fraudulent transactions: Analyze the dataset to identify features that are highly indicative of fraudulent activities.

3. Implement a real-time or near-real-time fraud detection system: Deploy the trained model into a system capable of processing transactions in real-time or near-real-time, enabling swift detection and prevention of fraudulent activities.

Key Tasks:

1. Data Analysis: Analyze historical credit card transaction data to gain insights into the characteristics of legitimate and fraudulent transactions.

2. Data Preprocessing: Handle imbalanced classes, preprocess the data, and engineer relevant features to enhance the model's performance.

3. Model Development: Build a classification model for fraud detection using machine learning algorithms.

4. Model Evaluation: Evaluate the performance of the developed model using appropriate metrics and interpret the importance of features in detecting fraudulent transactions.

5. System Implementation: Implement the trained model into a real-time or near-real-time system for continuous monitoring and detection of fraudulent activities.

Goals:

1. Achieve precision and recall scores above 90%: Evaluate the model's performance based on precision and recall metrics to ensure high accuracy in detecting fraudulent transactions.

2. Implement a real-time fraud detection system with low latency: Deploy the model into a system capable of processing transactions with minimal delay to enable timely detection and prevention of fraud.

3.Minimize financial losses due to fraudulent transactions: By accurately identifying and preventing fraudulent activities, aim to minimize financial losses incurred by the institution.

Scope:

1. Data Analysis and Preprocessing: Analyze and preprocess historical transaction data to prepare it for model training.

2. Model Development and Evaluation: Develop and evaluate the performance of the fraud detection model using machine learning techniques.

3. System Implementation: Implement the trained model into a real-time or near-real-time system for continuous monitoring and detection of fraudulent activities.

4. Continuous Monitoring and Improvement: Continuously monitor the performance of the system and incorporate improvements to enhance its effectiveness in detecting and preventing fraudulent transactions.

Implementation Code Summary:

1. Data Loading and Exploration: Load the credit card transaction dataset and perform initial exploration to understand its structure and contents.

2. Data Preprocessing: Handle missing values, analyze class distribution, and prepare the data for model training.

3. Under-Sampling: Perform under-sampling to balance the dataset by creating a sample dataset containing a similar distribution of normal and fraudulent transactions.

4. Model Training: Train a logistic regression model using the balanced dataset.

5. Model Evaluation: Evaluate the trained model's performance on both training and test datasets.

Conclusion:

In conclusion, the implementation of a fraud detection system using machine learning techniques is crucial for financial institutions to safeguard their assets and maintain trust with customers. By following the outlined objectives, tasks, and goals, the institution can develop an effective fraud detection system capable of accurately identifying and preventing fraudulent activities in real-time or near-real-time. The achieved accuracy on the training data of 98.29% and on the test data of 99.15% indicates the robustness of the implemented model in accurately distinguishing between legitimate and fraudulent transactions, thereby minimizing the risk of financial losses due to fraudulent activities.

