Itgalpura, Rajanukunte, Bengaluru - 560064

**School of Information Science**

A Project Report on

# CREDIT CARD FRAUD DETECTION USING PYTHON

|  |  |
| --- | --- |
| RAVI CHANDANA | 20221BCA0257 |
| CHANDRIKA | 20221BCA0220 |
| LIKHITHA | 20221BCA0221 |
| GEETHIKA REDDY | 20221BCA0219 |
| AISHWARYA | 20221BCA0215 |

Submitted by

Under the supervision of

**Guide name: MR. VIVEK BONGALE**

**Designation: Assistant Professor**

**Department: Department of Computer Science Engineering**

**ABSTRACT:**

With the increasing use of credit cards for online transactions, the risk of fraudulent activities has become a significant concern for financial institutions and consumers alike. In response to this challenge, this project proposes a novel approach for detecting credit card fraud using Python programming language.The project utilizes a dataset containing historical credit card transactions, including both legitimate and fraudulent ones. Through exploratory data analysis (EDA), various features such as transaction amount, time, and type are analyzed to identify patterns indicative of fraudulent behavior.Machine learning algorithms, including but not limited to logistic regression, decision trees, random forests are employed to develop predictive models capable of distinguishing between legitimate and fraudulent transactions. The proposed Python-based credit card fraud detection system offers a scalable and efficient solution to mitigate the risks.

**INTRODUCTION**

The widespread adoption of credit cards for online transactions has revolutionized the way we conduct financial transactions. However, this convenience comes with a significant challenge: the risk of credit card fraud. Fraudulent activities such as unauthorized transactions, identity theft, and account takeover pose a substantial threat to both financial institutions and consumers.

In response to this challenge, the development of robust fraud detection systems has become imperative. These systems leverage advanced technologies such as machine learning and data analysis to identify fraudulent transactions in real-time, thereby safeguarding the integrity of the financial ecosystem.

The project will begin with a comprehensive exploration of a dataset containing historical credit card transactions, encompassing both legitimate and fraudulent instances. Through exploratory data analysis (EDA), we will delve into the characteristics and patterns underlying fraudulent transactions, gaining insights that will inform the development of our detection algorithms.

Subsequently, we will employ a variety of machine learning algorithms, including logistic regression, decision trees, random forests to build predictive models capable of discerning between legitimate and fraudulent transactions. These models will be trained on labeled data and fine-tuned using techniques such as cross-validation and hyperparameter optimization to maximize their performance metrics.

## **Problem Statement:**

The Credit Card Fraud Detection Problem includes modeling past credit card transactions with the knowledge of the ones that turned out to be a fraud. This model is then used to identify whether a new transaction is fraudulent or not. Our aim here is to detect 100% of the fraudulent transactions while minimizing the incorrect fraud classifications.

**Goal Of** **The Project:**

This project aims to contribute to the ongoing efforts in credit card fraud detection by proposing a Python-based approach that combines data analysis, machine learning, and optimization techniques. By harnessing the power of Python's rich ecosystem of libraries and frameworks, we seek to develop a scalable and efficient solution for detecting fraudulent activities in credit card transactions.

**TECHNIQUES:**

**1.Supervised Learning Algorithms.**

**2.Unsupervised Learning Algorithms.**

**3.Feature Engineering.**

**4.Regression Algorithm**

**ATTRIBUTES:**

1. **Transaction Amount**: The amount of money involved in the transaction, which can vary widely between legitimate and fraudulent transactions.
2. **Transaction Time**: The timestamp indicating when the transaction occurred. Patterns in transaction times may reveal anomalies that could indicate fraudulent activity.
3. **Merchant Category Code (MCC)**: A four-digit code assigned to merchants to classify the type of business they operate. Certain MCCs may be associated with a higher risk of fraudulent transactions.
4. **Transaction Type**: Whether the transaction is a purchase, cash advance, transfer, or other types of transactions. Fraudulent transactions may exhibit distinct patterns in transaction types.
5. **Cardholder Verification Method**: The method used to verify the identity of the cardholder, such as PIN verification, signature verification, or no verification.
6. **Card Type**: The type of credit card used for the transaction (e.g., Visa, Mastercard, American Express).
7. **Distance from Cardholder's Home**: The geographical distance between the location of the transaction and the cardholder's registered home address. Unusual distances may indicate potential fraud.
8. **Number of Transactions in a Time Window**: The frequency of transactions within a specified time window (e.g., number of transactions in the last hour). Rapidly occurring transactions may be indicative of fraudulent behavior.
9. **Transaction Status**: Whether the transaction was approved, declined, or pending authorization.
10. **Account Age**: The length of time the account associated with the credit card has been active. New accounts or accounts with minimal activity may be more susceptible to fraud.
11. **Previous Fraud History**: Whether the cardholder or the associated account has a history of previous fraudulent activity.

**References:**

Journal: https://ieeexplore.ieee.org/

Dataset: <https://www.kaggle.com/>