Credit Card Fraud Detection using Machine Learning

Stage II
Project Review

Project Guide – Prof. Ashwini Bhosale

Presenting –

- Kaveri Diwanji B190304252
- Samarth Malegaonkar B190304263
- Samrajya Pujari B190304291
- Samreen Shaikh B190304304

Credit Card Fraud Detection using Machine Learning

Problem Statement

With businesses moving online, fraud and abuse in online systems are constantly increasing as well. Traditionally, rule-based fraud detection systems are used to combat online fraud, but these rely on a static set of rules created by human experts. This project uses machine learning to create models for fraud detection that are dynamic, self-improving, and maintainable. Importantly, they can scale with the online business.

Study by RBI

The average daily transactions is 100 million now for a volume of Rs 5 trillion.

Payments through digital modes are expected to jump to 1.5 billion transactions, worth Rs 15 trillion a day in five years, the Reserve Bank of India (RBI) estimates.

Statistics

2020 - 389,845 Credit card fraud reports.
2021 - 1,862 data breaches
2022 - banking frauds in India amounted for 1.38 trillion Indian rupees

Main Objective of Proposed Work

Automate the detection of potentially fraudulent activity, and the flagging of that activity for review

overview

The project spans over different key stages.

These stages include **data collection, preprocessing, model development**, **feature engineering** and **selection**, **Streamlit** application development, and deployment.

The use of machine learning algorithms will enable accurate fraud detection, while the Streamlit platform will ensure a user-friendly experience.

Data Collection and Preprocessing

Data collection involves acquiring a dataset containing credit card transactions, including both fraudulent and legitimate ones.

Preprocessing the data is crucial to ensure its quality and reliability.

Challenges such as class imbalance, missing values, and duplicates need to be addressed through appropriate techniques.

Model Development

Machine learning algorithms play a crucial role in detecting patterns and anomalies within the credit card transaction data.

logistic regression, decision trees, random forests, and gradient boosting, are explored and evaluated.

performance metrics and their ability to accurately detect fraud.

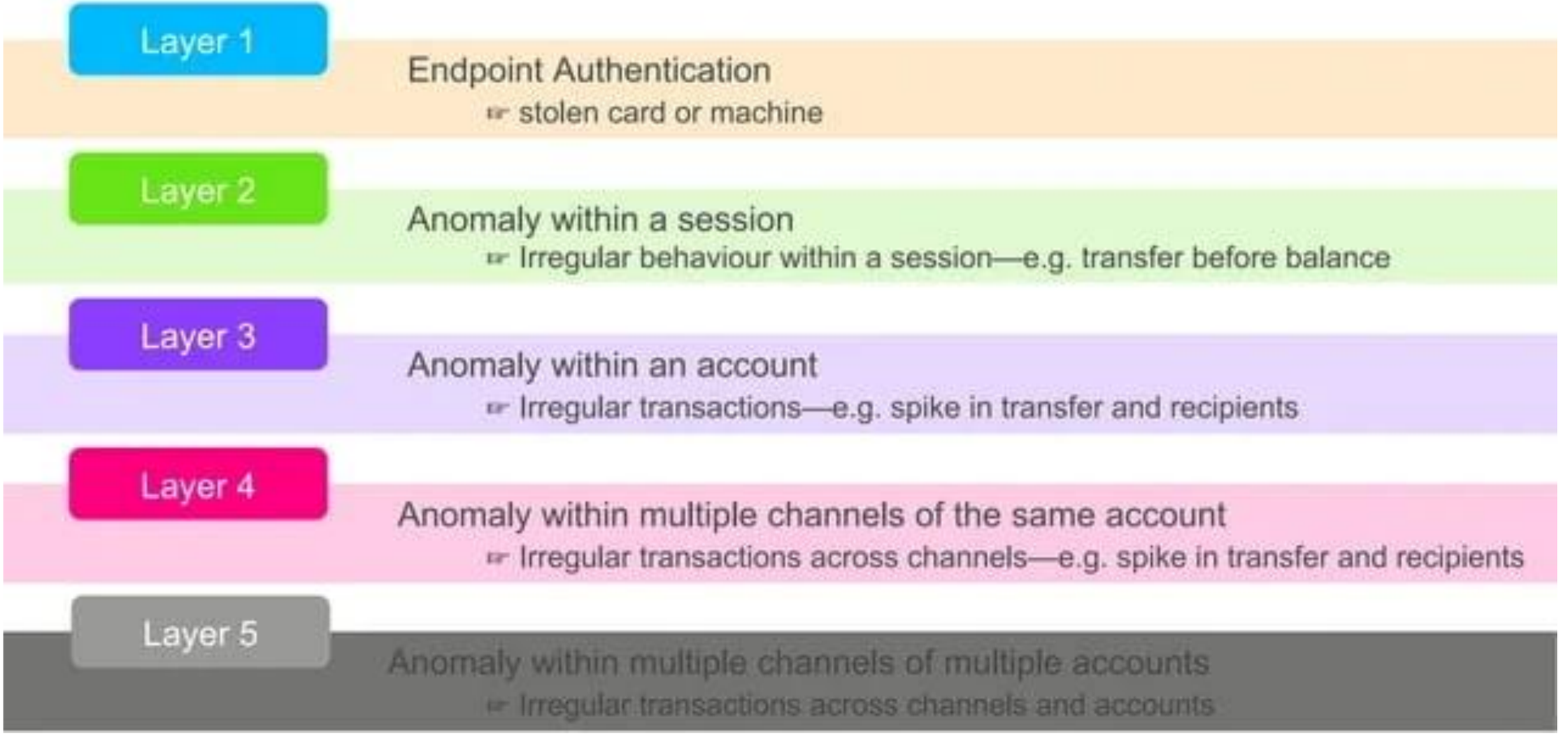
Feature Engineering and Selection

Feature engineering involves extracting meaningful features from the dataset that contribute to fraud detection.

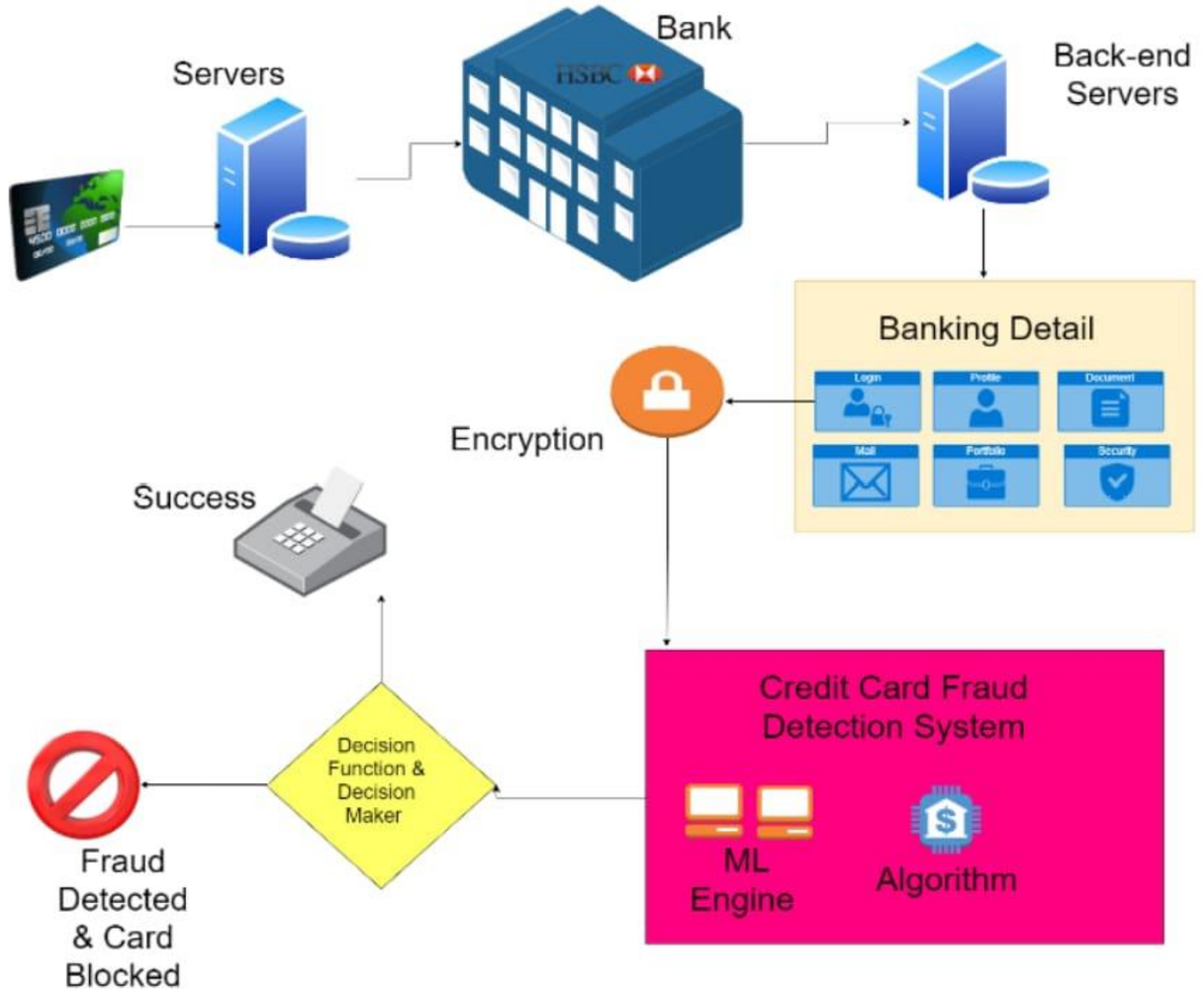
Feature selection techniques, such as Principal Component Analysis (PCA) or Recursive Feature Elimination (RFE), will be employed to identify the most relevant features.

These processes aim to enhance the model's performance and improve its ability to differentiate between fraudulent and legitimate transactions.

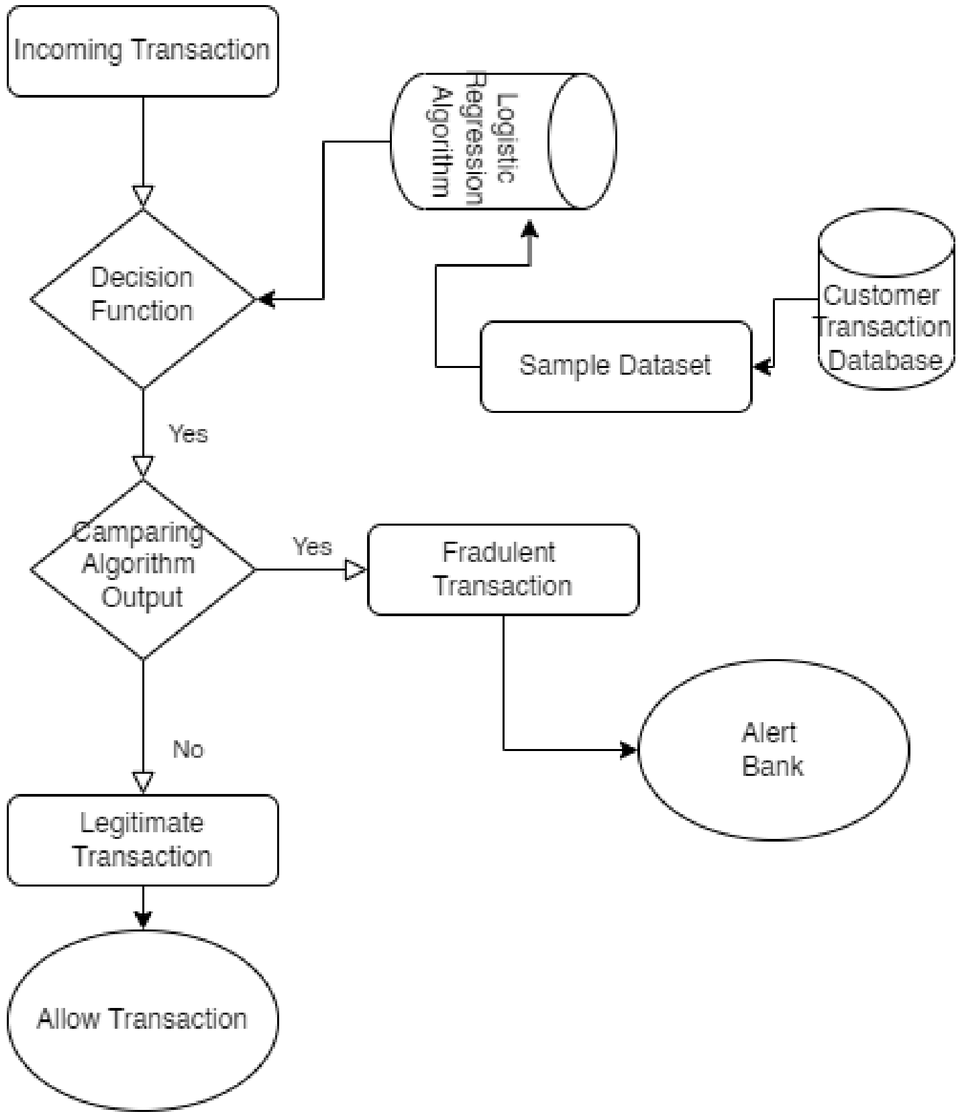
Types of Fraudulent Behaviour



Proposed Work Flow Chart



Flow Chart


Incoming Transaction
Logistic Regression Algorithm
Decision Function
Customer Transaction Database
Sample Dataset
Yes
Camparing Algorithm Output
Yes
Fradulent Transaction
Alert Bank
No
Legitimate Transaction
Allow Transaction

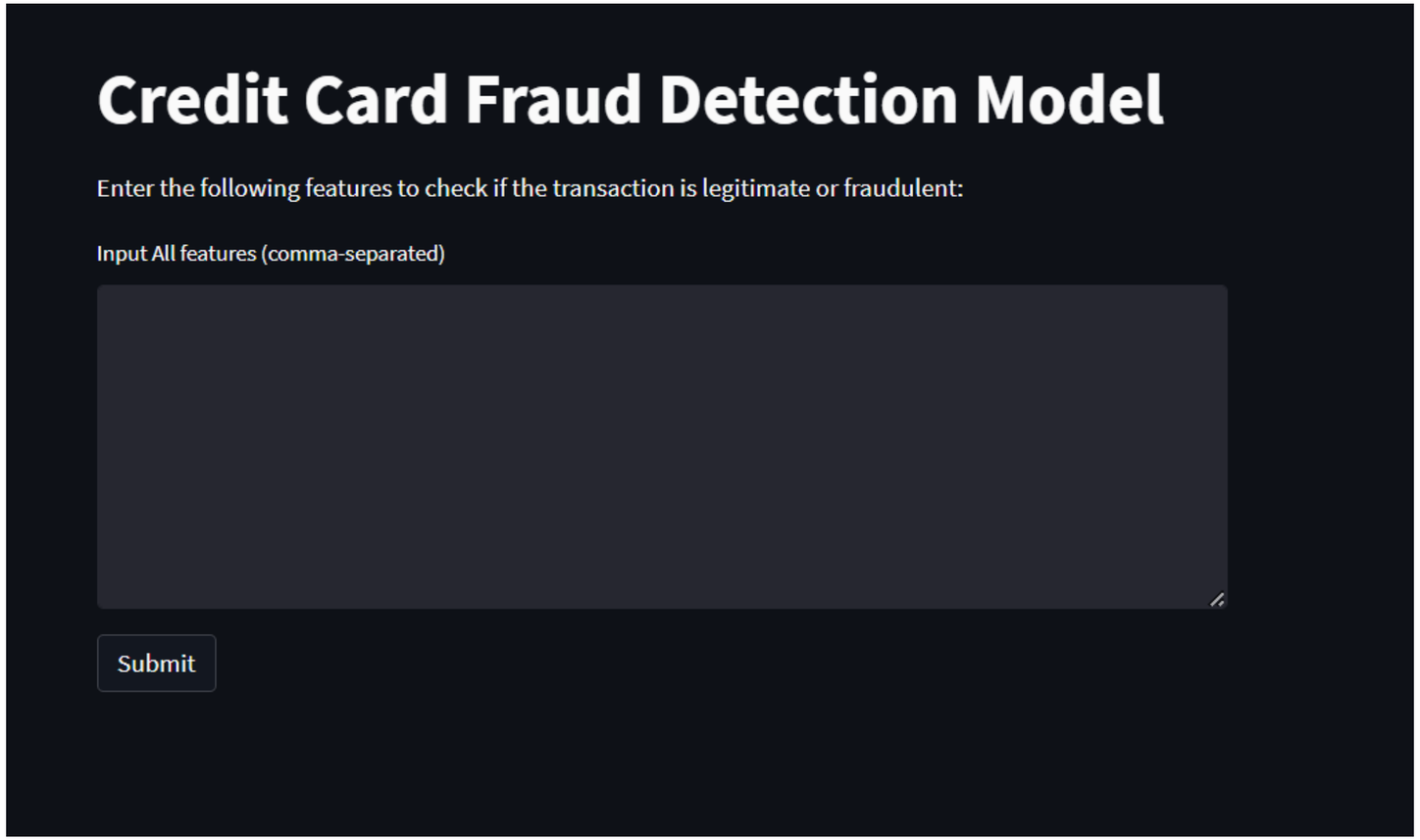
Streamlit Application Development

Streamlit is a powerful platform for building interactive web applications.

The credit card fraud detection application has a user-friendly interface, allowing users to input transaction details.

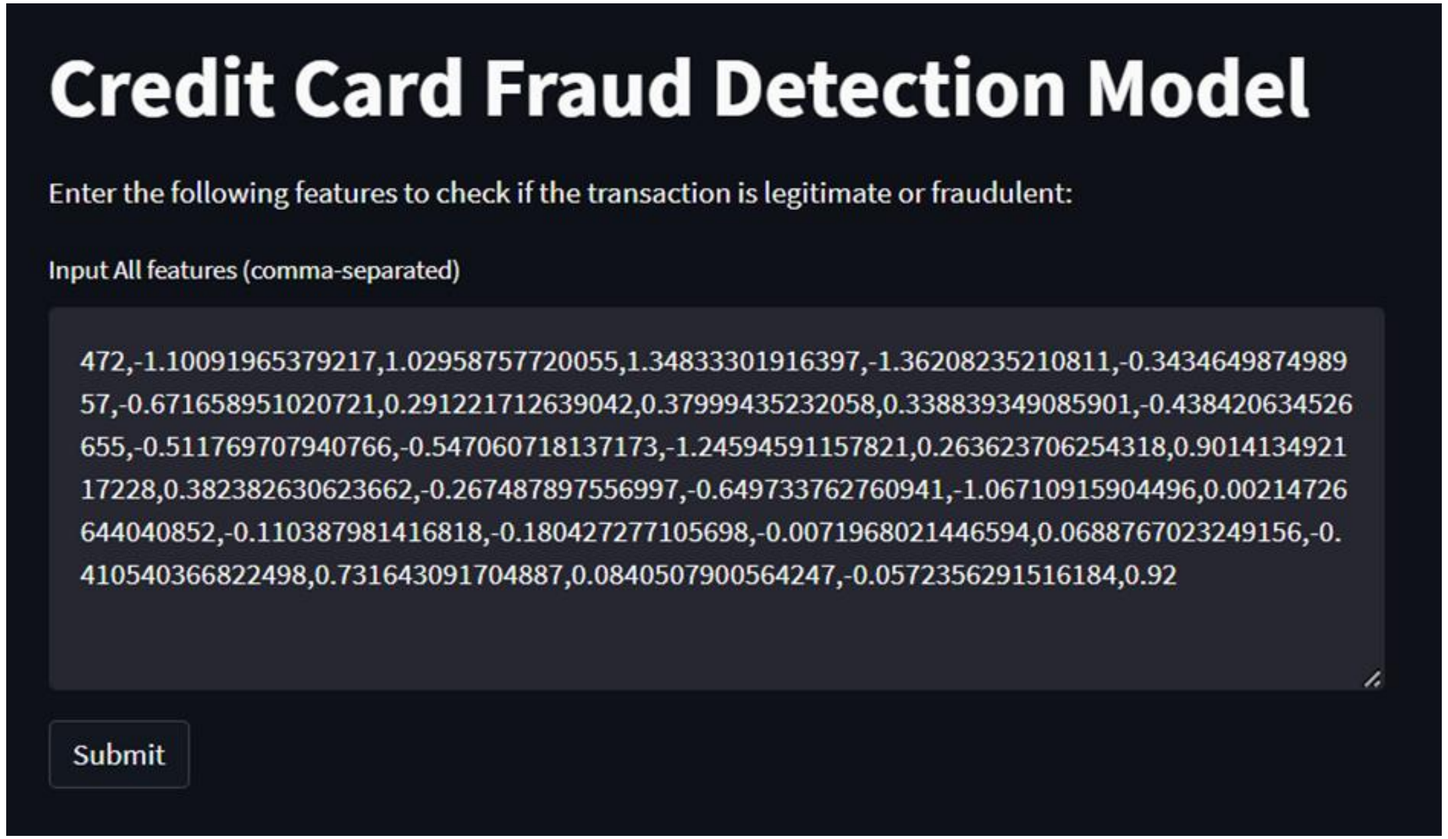
The application process the inputs using the deployed machine learning model and display real-time fraud detection results.

How Application work?



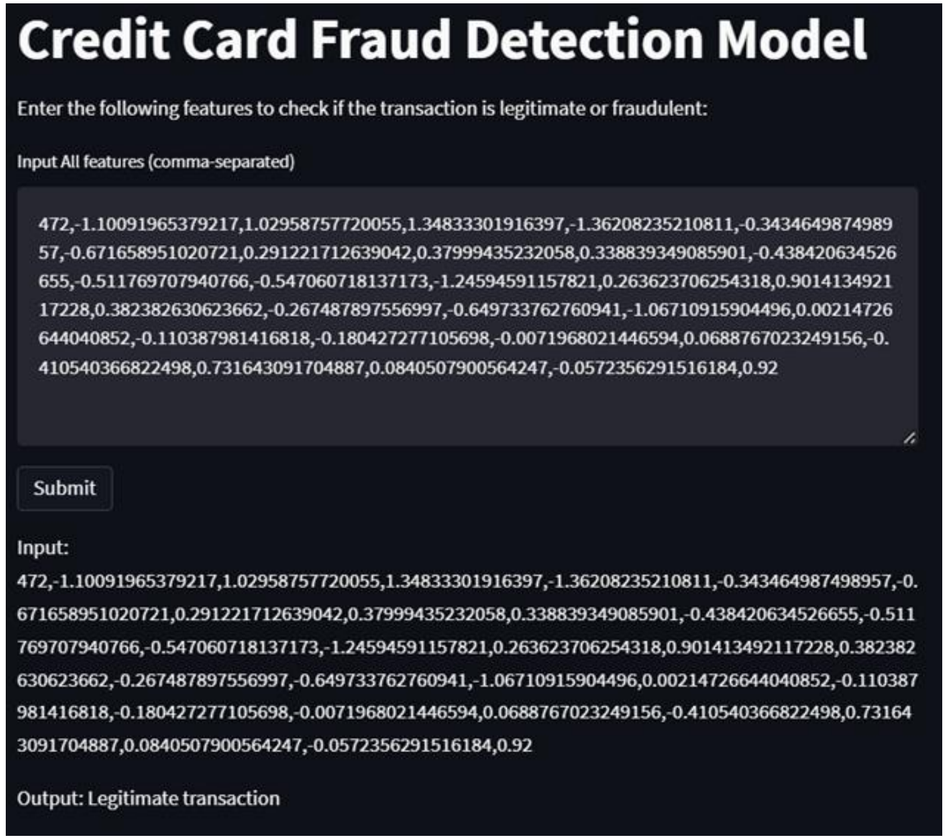
User Interface of the Application

How Application work?



Transaction details

How Application work?



Output

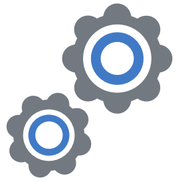
Future Scope

- **Incorporating Advanced ML Techniques**
- **Continuous Model Training**
- **Integration with External Data Sources**
- **Collaboration with Industry Stakeholders**
- **Training and Deployment using AWS SageMaker**

What we Used ?

Software Requirements

- Operating system – Windows 10/11
- Visual Studio Community - 2022
- Microsoft Office 2019
- Any Web Browser – Latest Version
- Python
- Jupyter Notebook
- Streamlit



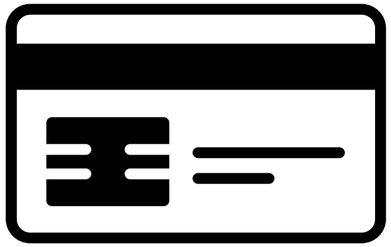
Hardware Requirements

- Processor - 64-bit, four-core, 2.5 GHz minimum per core
- RAM– 8 GB Minimum
- Hard Disk – 256 GB Required

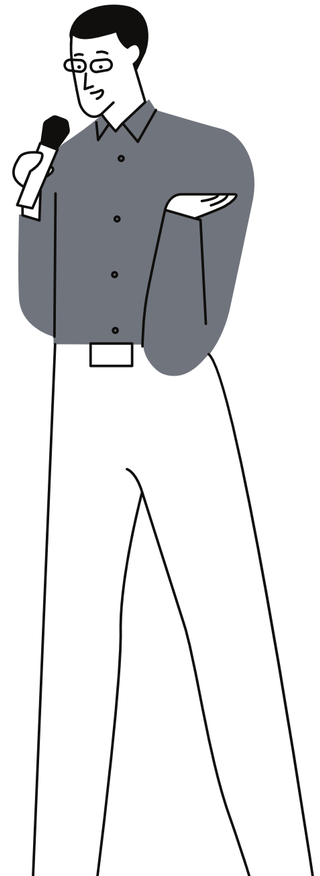
Work Carried Out from stage I

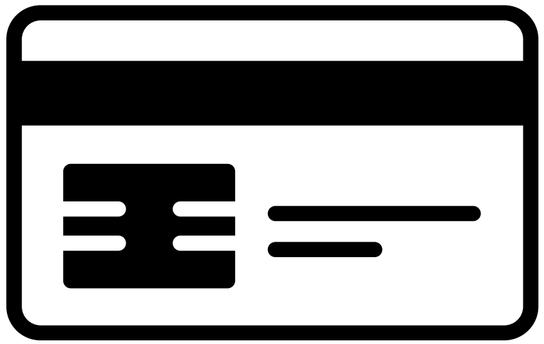
Application Development and Deployment

explore new Algorithms to detect Fraudulent Transactions



Open for Suggestion!





Thank You!