ANNAMACHARYA INSTITUTE OF TECHNOLOGY & SCIENCES
UTUKUR, C. K. DINNE (V&M), KADAPA, YSR DIST.
Approved by AICTE, New Delhi & Affiliation to JNTUA, Anantapuramu.
Accredited by NAAC with 'A' Grade, Bangalore.

DETECTING FRAUDLENT CREDIT CARD TRANSACTIONS USING ENSEMBLE METHODS

submitted by

P. Swathi	(21HM1A05A4)
K. Ajith Kumar	(21HM1A0574)
S. Shaikshavali	(22HM5A0510)
P. Jyoshna	(21HM1A05A7)

Under the Esteemed Guidance of

Miss T. Anusha , M.tech.,
Assistant professor,
CSE Department

CONTENTS

- Abstract
- Introduction
- Existing system
- Disadvantages
- Proposed system
- Advantages
- System requirements
- Algorithms
- Architecture
- Modules
- UML Diagrams
- Testing
- Screenshots
- Conclusion
- Future work
- References

ABSTRACT

Detecting fraudulent credit card transactions is a major challenge for banks. Hackers can impersonate cardholders and conduct fraud. It explores ensemble learning methods to detect fraud using the Sparkov synthetic dataset and a real european consumer dataset. XGBoost, random forests, and naive Bayes classifiers are applied and evaluated. Performance is measured using accuracy, precision, recall, and F1 score.

INTRODUCTION

A credit card is a financial tool that allows users to make purchases, pay bills, and withdraw cash on credit. It is widely used for shopping, online transactions, travel bookings, bill payments, and emergency expenses.

EXISTING SYSTEM

Banks use statistical methods like probability distributions and box plots to detect unusual transactions. Decision trees and random forests help classify transactions as fraud or genuine. However, these methods rely on known fraud patterns, making it difficult to detect new fraud techniques. It also struggles with real-time fraud detection and have a high false positive rate, leading to many legitimate transactions being flagged as fraud.

DISADVANTAGES

- Low Accuracy
- Complexity of Data
- Data Requirement
- Incorrect Labeling

PROPOSED SYSTEM

Ensemble methods can be used to learn from both real and synthetic datasets. It used to detect fraud using the Sparkov synthetic dataset and a real european consumer dataset. Instead of relying on predefined fraud patterns, it identifies new threats using XGBoost and Random Forest for classification, Autoencoders and Isolation Forest for anomaly detection. This system ensures real-time transaction monitoring and strengthens security.

ADVANTAGES

- Improved fraud detection
- Real-time fraud prevention
- Reduced financial losses
- Enhanced transaction security
- Adaptive to new fraud techniques

SYSTEM REQUIREMENTS

SOFTWARE REQUIREMENTS

➢ Operating system	:	Windows 7 Ultimate.
➢ Front-End	:	Python.
➢ Back-End	:	Django-ORM
➢ Designing	:	HTML, CSS, Javascript.
➢ Data Base	:	MySQL (WAMP Server).

HARDWARE SYSTEM CONFIGURATION

Processor	:	Pentium –IV
RAM	:	4 GB (min)
Hard Disk	:	20 GB
Key Board	:	Standard Windows Keyboard
Monitor	:	SVGA

ALGORITHMS

- **Naive Bayes**:

Naive Bayes is a probabilistic classification algorithm based on Bayes' Theorem, assuming that features are independent. It works well with small and large datasets.

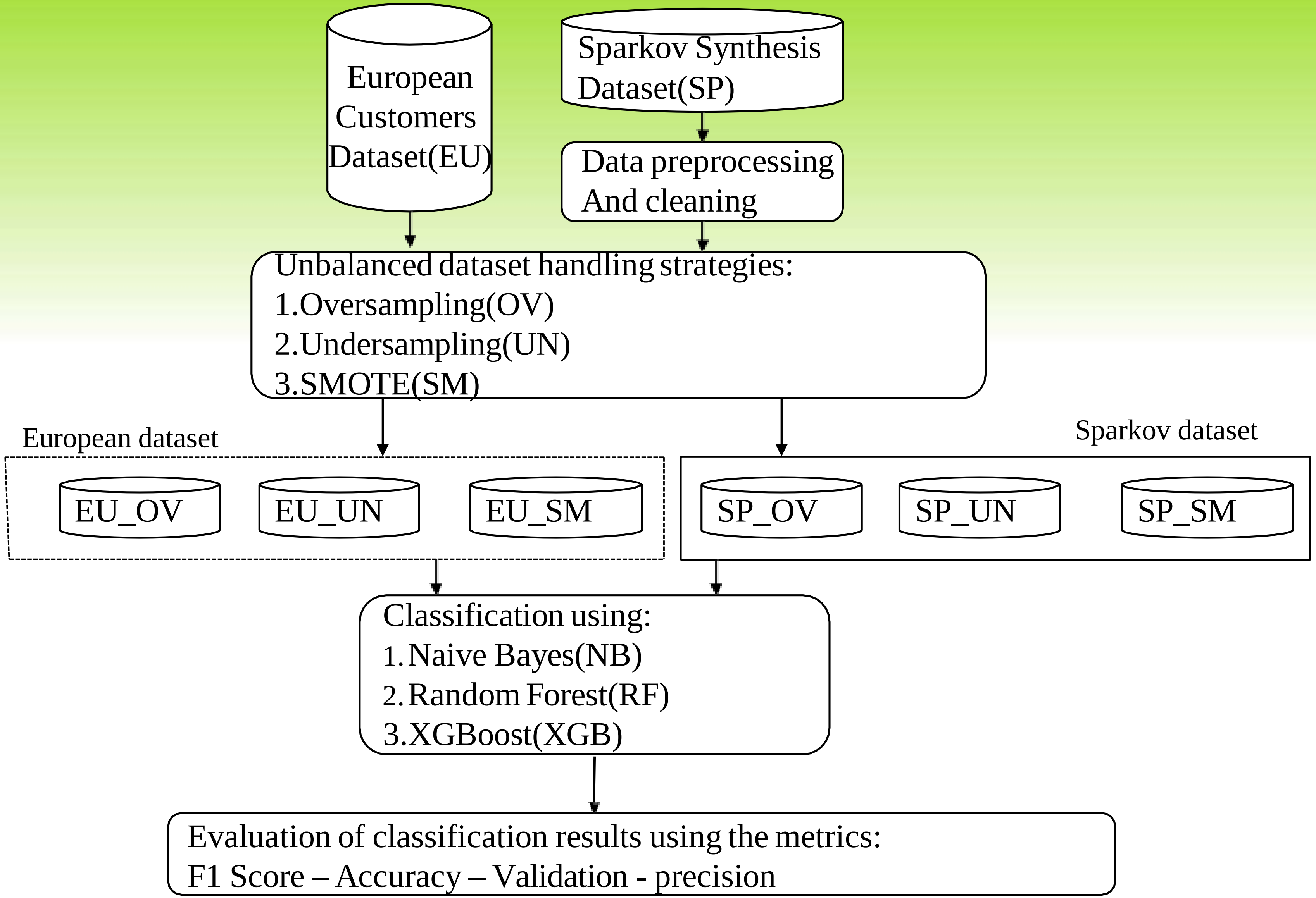
- **Random Forest:**

It's a type of ensemble learning method. A random forest (RF) is a collection of decision trees that are used to make predictions**.**

- **XGBoost:**

XGBoost is a robust machine-learning algorithm that can help you understand your data and make better decisions. XGBoost is an implementation of gradient-boosting decision trees.

ARCHITECTURE



European Customers Dataset(EU)

Sparkov Synthesis Dataset(SP)

Data preprocessing And cleaning

Unbalanced dataset handling strategies:
1.Oversampling(OV)
2.Undersampling(UN)
3.SMOTE(SM)

European dataset

EU_OV EU_UN EU_SM

Sparkov dataset

SP_OV SP_UN SP_SM

Classification using:
1.Naive Bayes(NB)
2.Random Forest(RF)
3.XGBoost(XGB)

Evaluation of classification results using the metrics:
F1 Score – Accuracy – Validation - precision

MODULES

USER MODULE

- Users log in to access the system.
- Upload datasets and apply algorithms for processing.

Operations:

- Login & Authentication
- Upload Dataset
- Apply Algorithm
- Predict Results
- View Results

SYSTEM MODULE

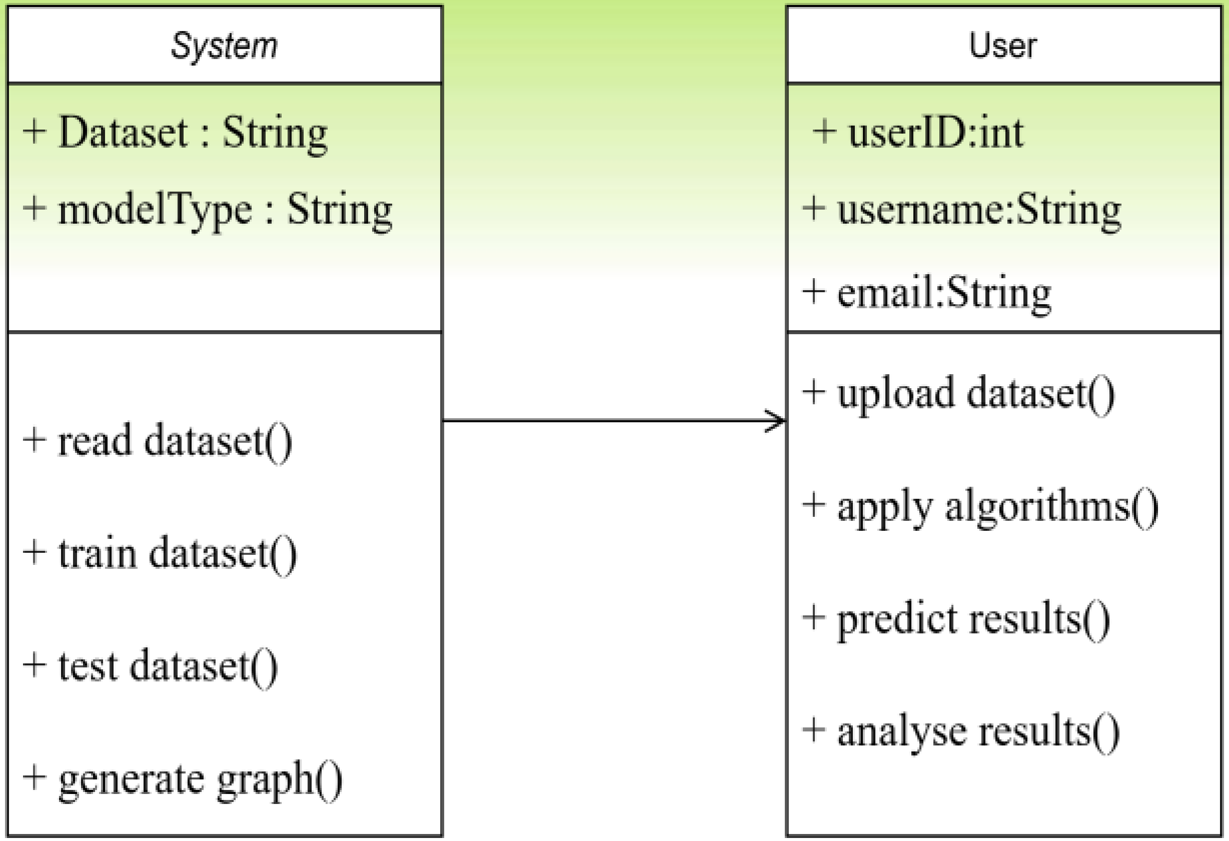
- Processes the dataset by training and testing it.
- Generates predictions and displays analysis results.

Operations:

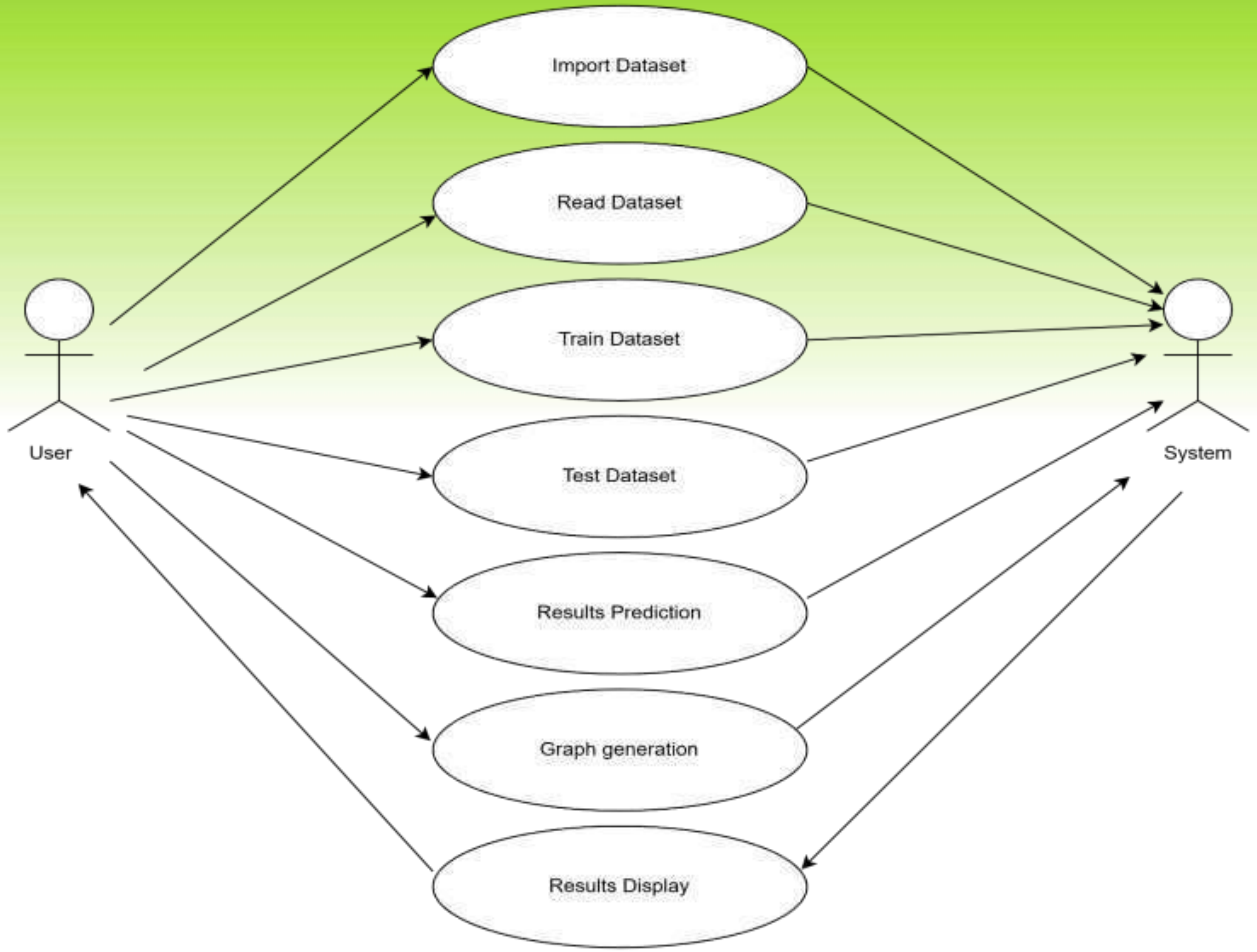
- Read Dataset
- Train Dataset
- Test Dataset
- Predict Results
- Display Results

UML DIAGRAMS

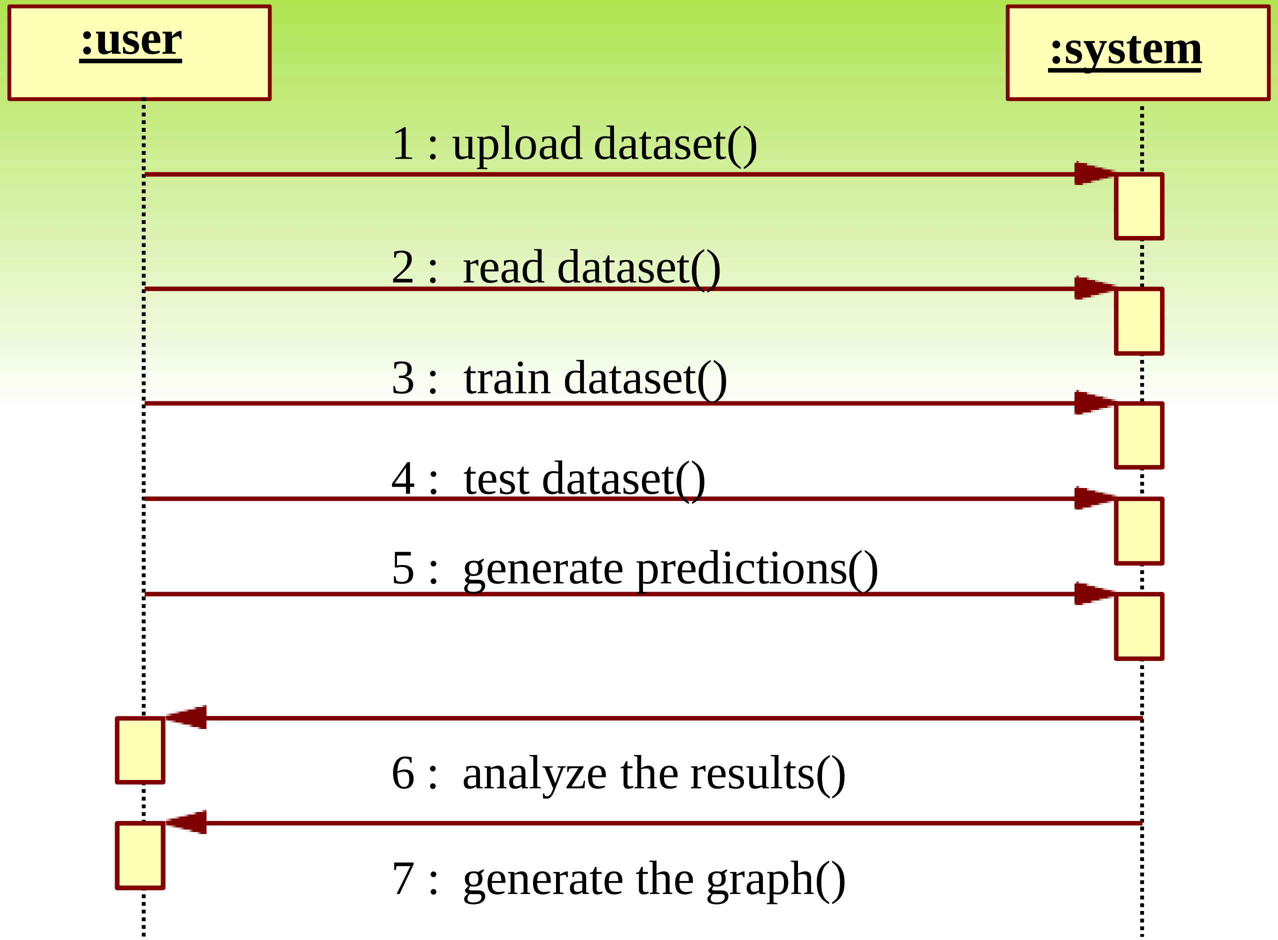
CLASS DIAGRAM



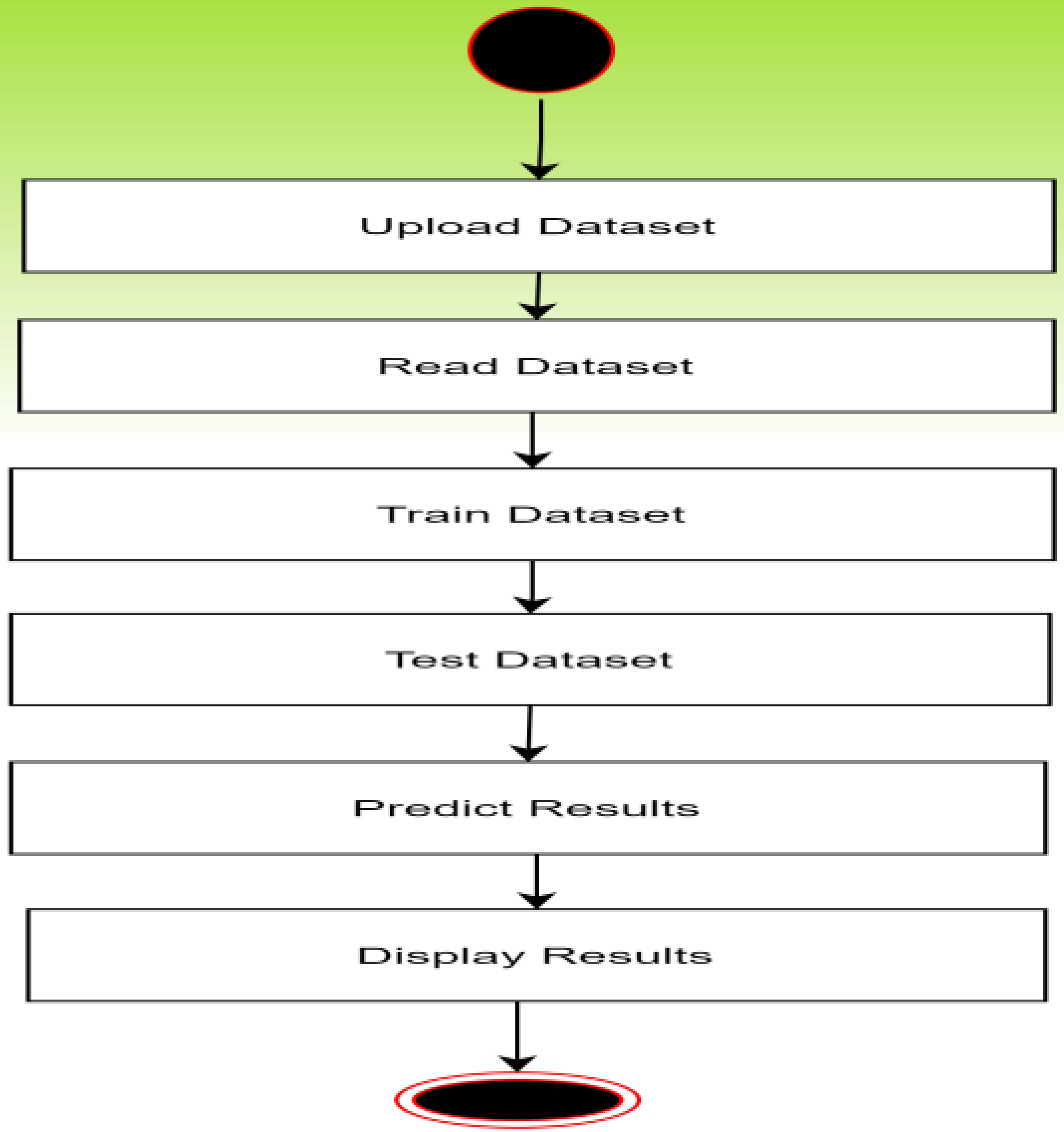
USE CASE DIAGRAM



SEQUENCE DIAGRAM



ACTIVITY DIAGRAM



Upload Dataset

Read Dataset

Train Dataset

Test Dataset

Predict Results

Display Results

TESTING

- **Unit Testing**

Verifies the correct functioning of individual components such as data preprocessing, feature extraction, and fraud classification.

- **Integration Testing**

Ensures seamless interaction between dataset handling, model training, and fraud prediction modules.

- **Performance Testing**

Evaluates accuracy, precision, recall, and F1 Score while testing efficiency on large datasets and real-time transactions.

- **Output Testing**

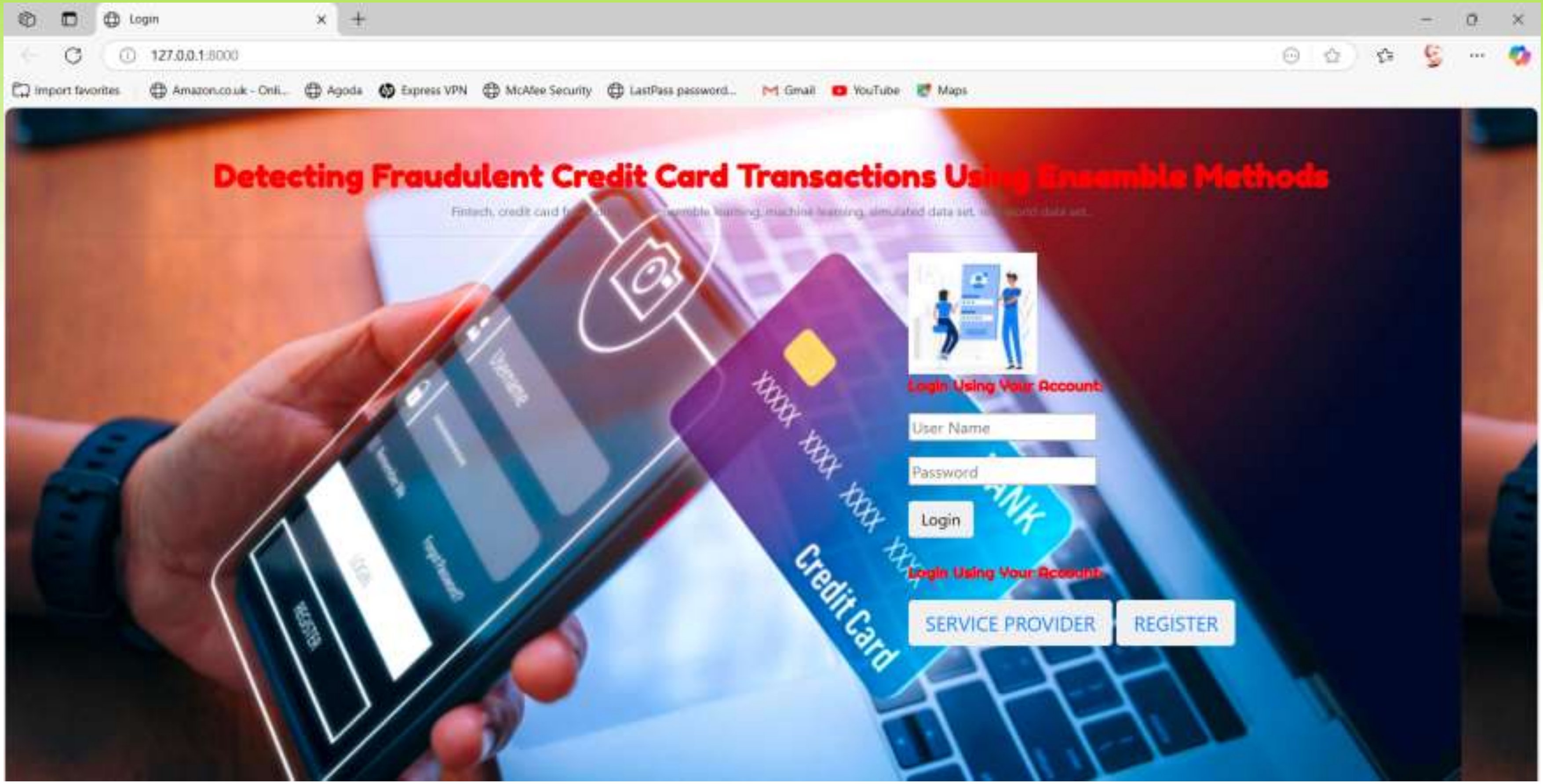
Checks if the model correctly classifies fraudulent and non-fraudulent transactions by comparing predictions with actual labels.

- **Regression Testing**

Ensures that model updates or optimizations do not introduce errors or degrade performance over time.

SCREENSHOTS

HOME



Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Fintech, credit card fraud detection, ensemble learning, machine learning, simulated data set, real world data set.

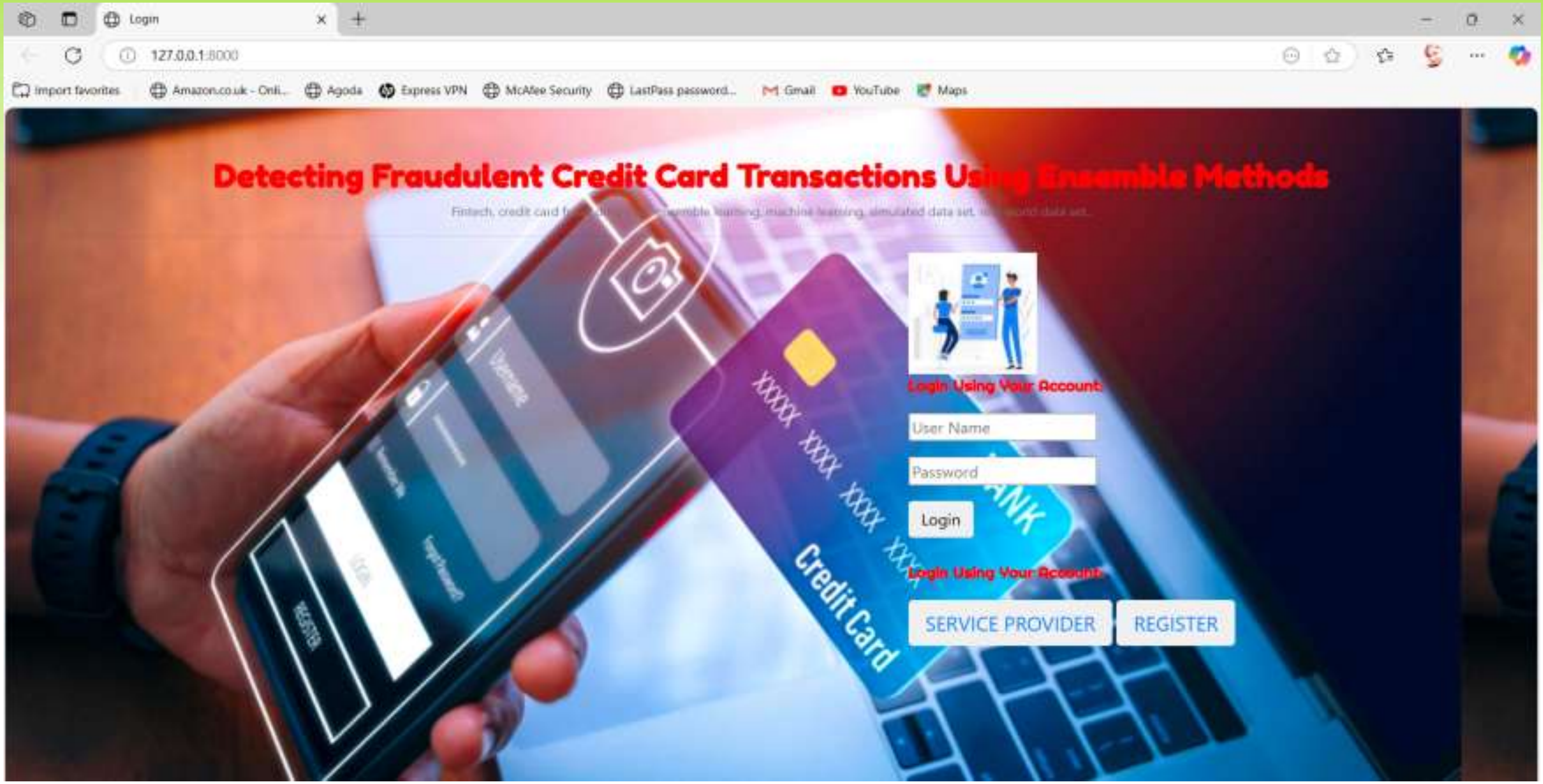
Login Using Your Account

User Name

Password

Login Using Your Account:

USER LOGIN



127.0.0.1:8000

Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Fintech, credit card fraud detection, ensemble learning, machine learning, simulated data set, real-world data set.

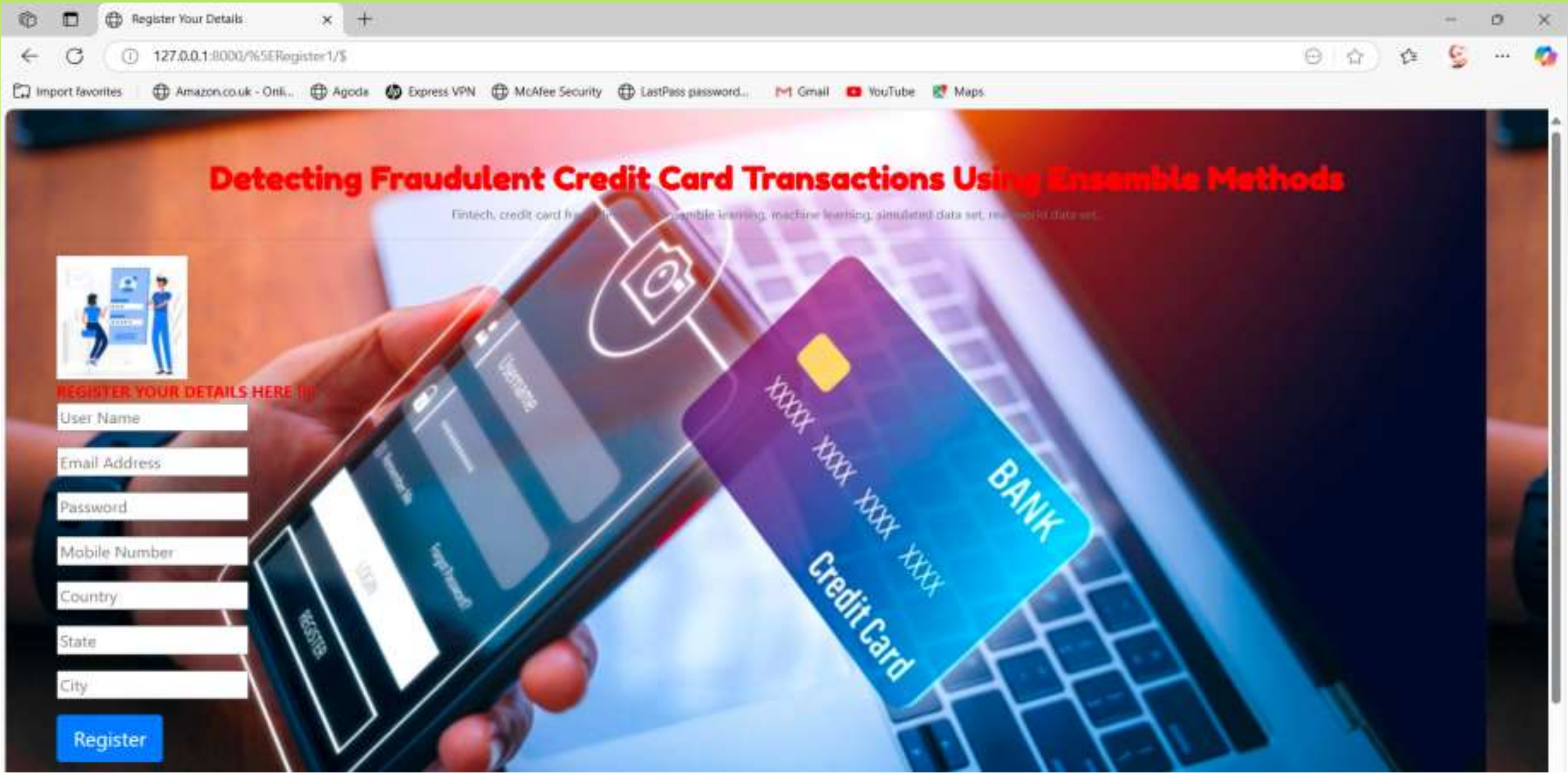
Login Using Your Account

User Name

Password

Login Using Your Account

USER REGISTRATION



Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Fintech, credit card fraud detection, ensemble learning, machine learning, simulated data set, real-world data set.

REGISTER YOUR DETAILS HERE !!!

User Name

Email Address

Password

Mobile Number

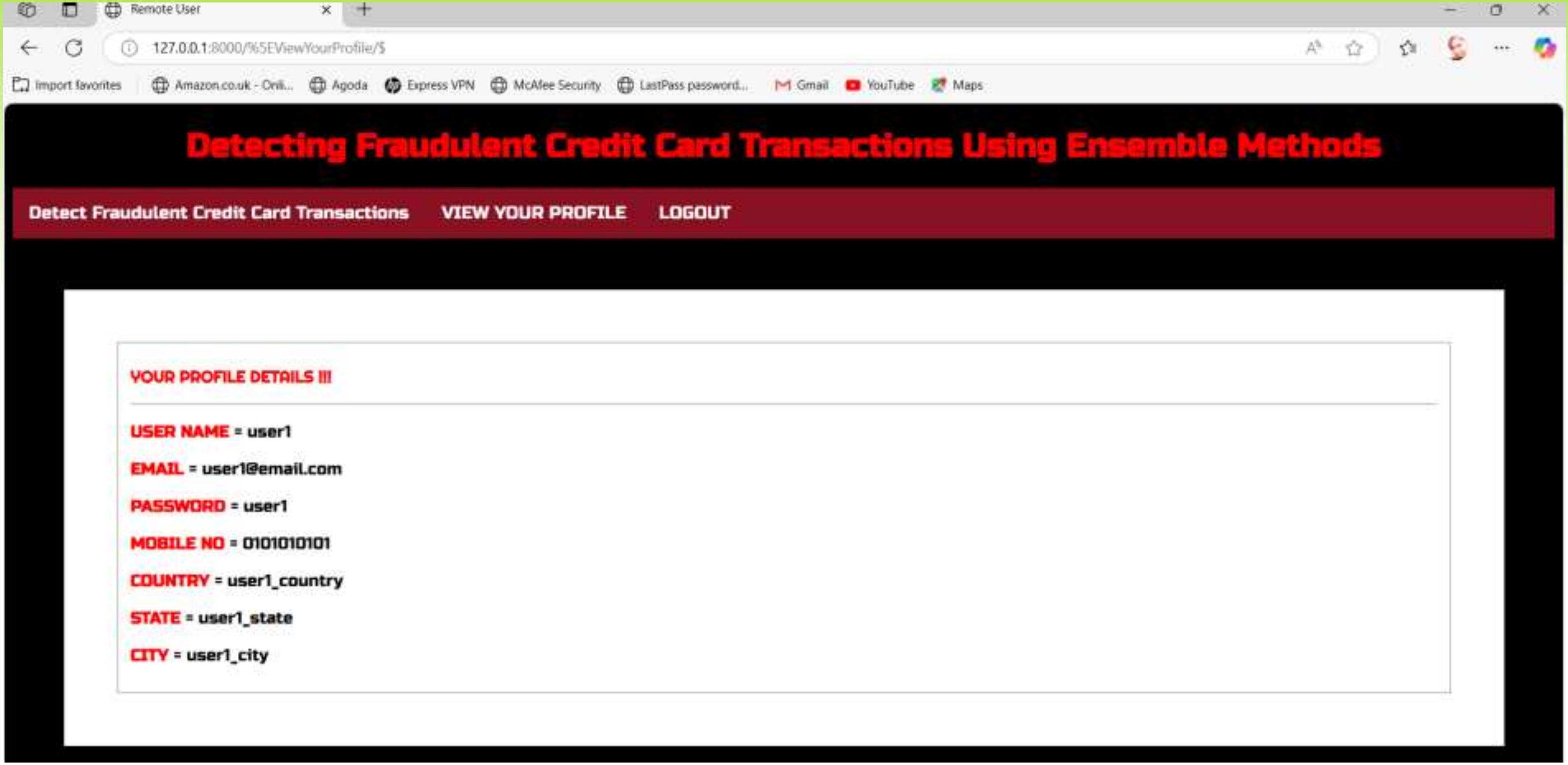
Country

State

City

Register

USER CREDENTIALS



Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

YOUR PROFILE DETAILS !!!

USER NAME = user1

EMAIL = user1@email.com

PASSWORD = user1

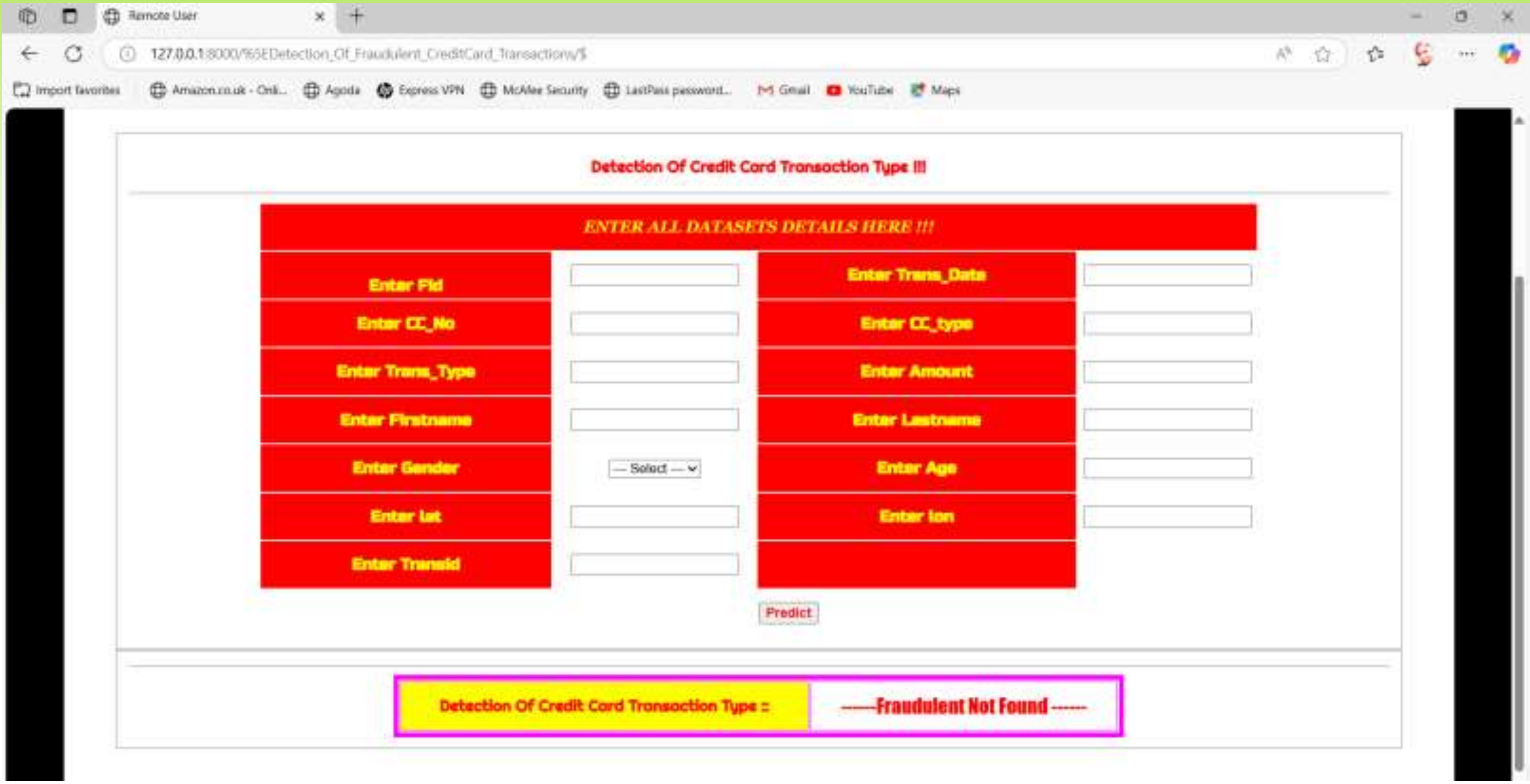
MOBILE NO = 0101010101

COUNTRY = user1_country

STATE = user1_state

CITY = user1_city

PREDICTION OF FRAUD



Detection Of Credit Card Transaction Type !!!

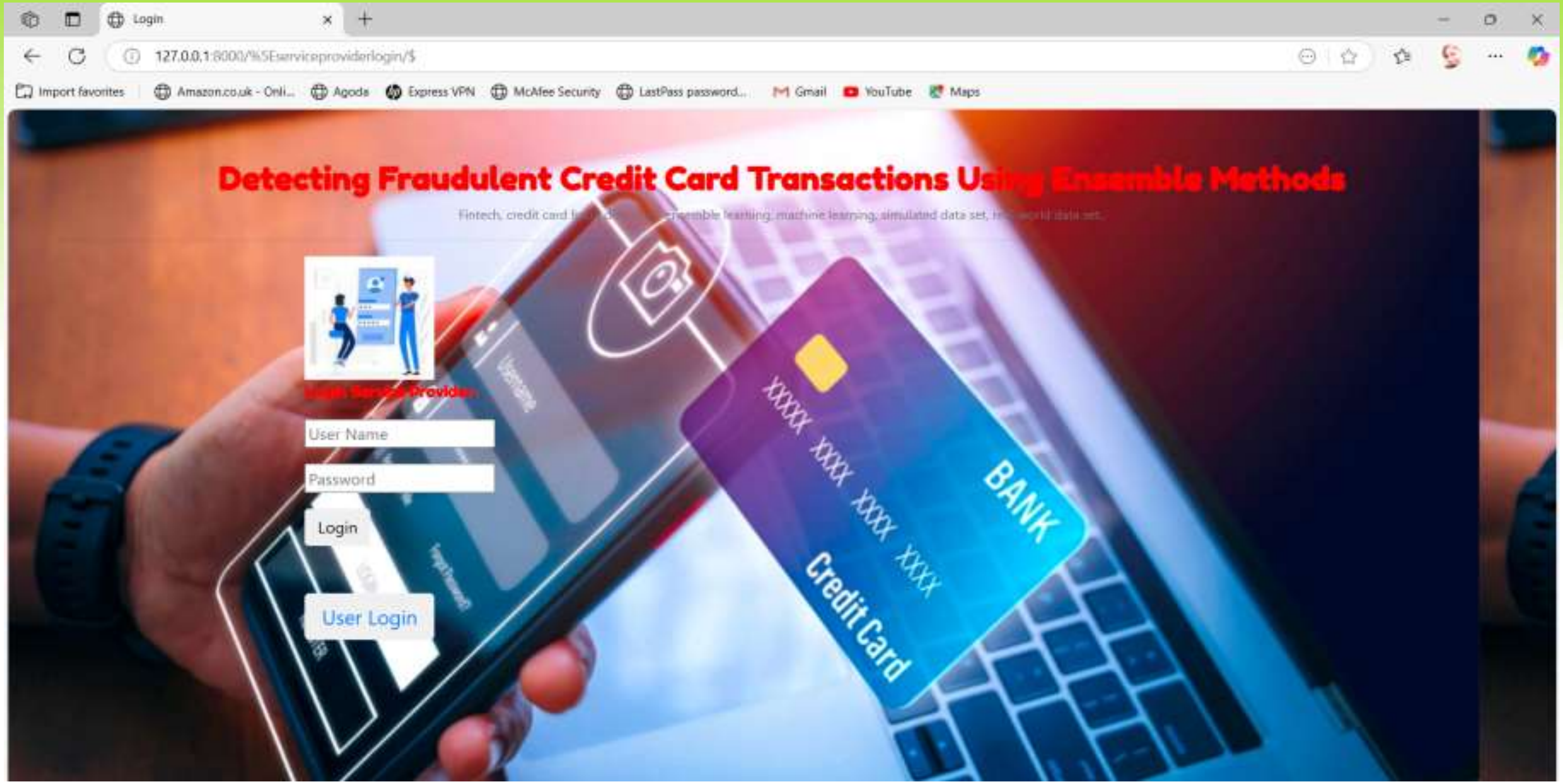
ENTER ALL DATASETS DETAILS HERE !!!

Enter Fid		Enter Trans_Date	
Enter CC_No		Enter CC_type	
Enter Trans_Type		Enter Amount	
Enter Firstname		Enter Lastname	
Enter Gender	-- Select --	Enter Age	
Enter lat		Enter lon	
Enter Transid			

Predict

Detection Of Credit Card Transaction Type ::	------Fraudulent Not Found ------

ADMIN LOGIN



Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Fintech, credit card fraud detection, ensemble learning, machine learning, simulated data set, real-world data set.

Login Service Provider

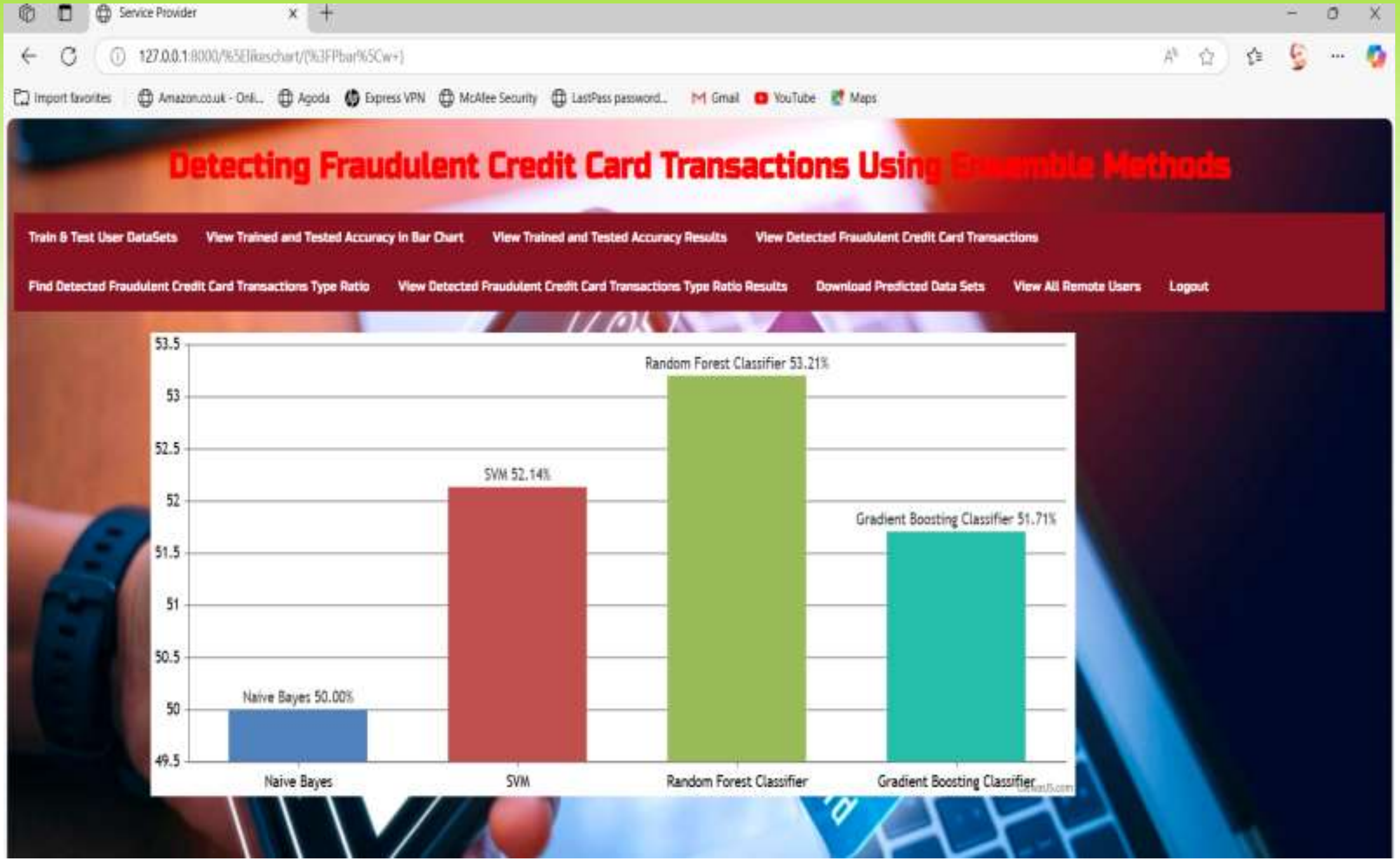
User Name

Password

Login

User Login

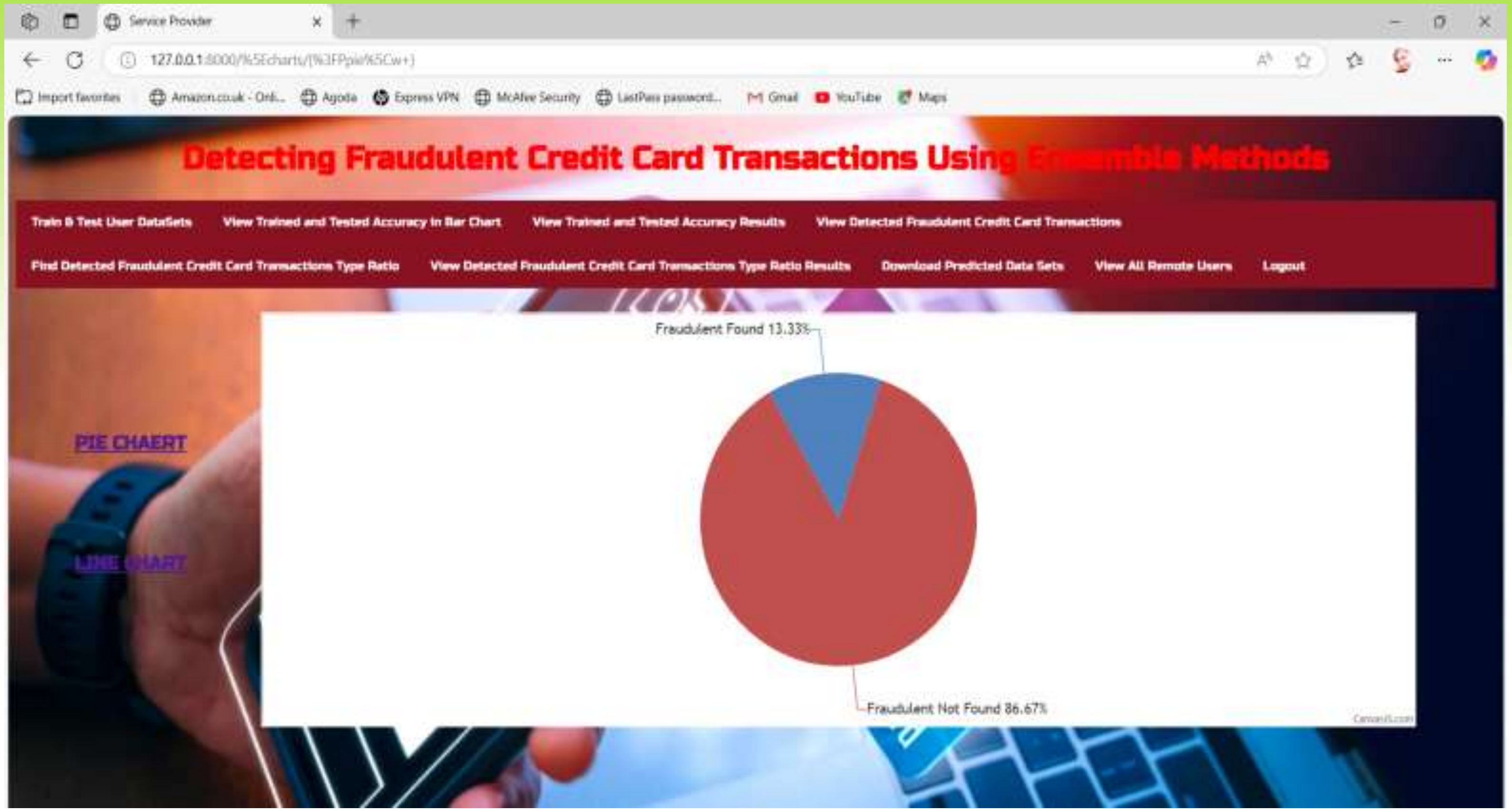
ACCURACY RESULTS BAR CHART



Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Service Provider
127.0.0.1:8000/%5Elikeschart/(%3FPbar%5Cw+)
Import favorites
Amazon.co.uk - Onli...
Agoda
Express VPN
McAfee Security
LastPass password...
Gmail
YouTube
Maps
Train & Test User DataSets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Detected Fraudulent Credit Card Transactions
Find Detected Fraudulent Credit Card Transactions Type Ratio
View Detected Fraudulent Credit Card Transactions Type Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout
53.5
53
52.5
52
51.5
51
50.5
50
49.5
Naive Bayes 50.00%
SVM 52.14%
Random Forest Classifier 53.21%
Gradient Boosting Classifier 51.71%
Naive Bayes
SVM
Random Forest Classifier
Gradient Boosting Classifier
CanvasJS.com

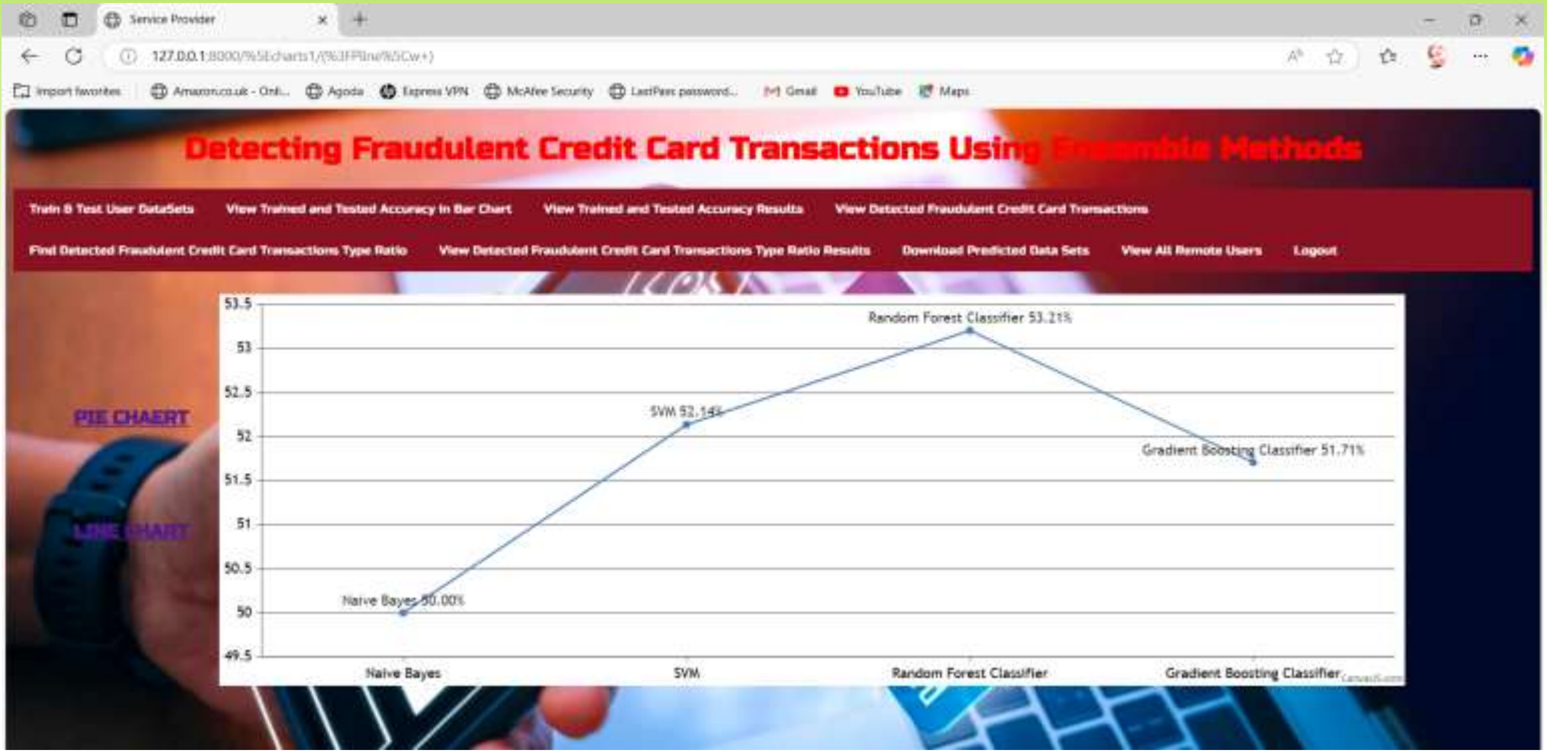
PIE CHART



Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Service Provider
127.0.0.1:8000/%5Echarts/(%3FPpie%5Cw+)
Import favorites
Amazon.co.uk - Onli...
Agoda
Express VPN
McAfee Security
LastPass password...
Gmail
YouTube
Maps
Train & Test User DataSets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Detected Fraudulent Credit Card Transactions
Find Detected Fraudulent Credit Card Transactions Type Ratio
View Detected Fraudulent Credit Card Transactions Type Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout
PIE CHAERT
LINE CHART
Fraudulent Found 13.33%
Fraudulent Not Found 86.67%
CanvasJS.com

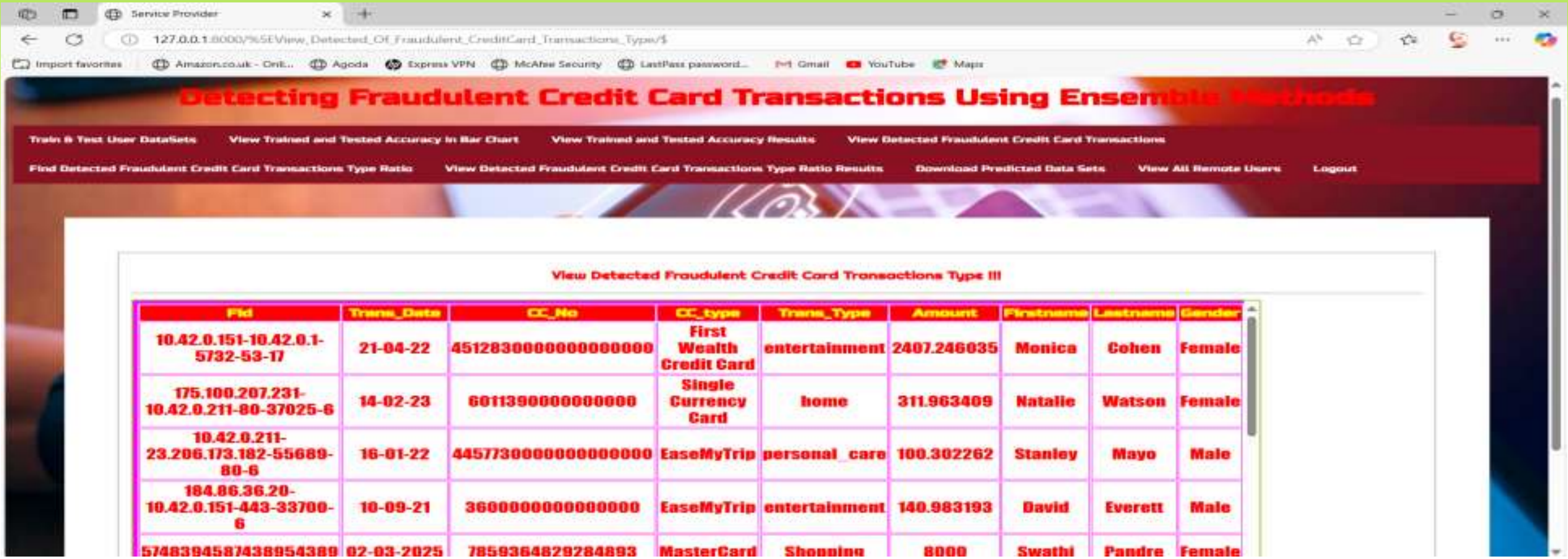
LINE CHART



Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Train & Test User DataSets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Detected Fraudulent Credit Card Transactions
Find Detected Fraudulent Credit Card Transactions Type Ratio
View Detected Fraudulent Credit Card Transactions Type Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout

FRAUD DETECTION DETAILS



Service Provider
127.0.0.1:8000/%5EView_Detected_Of_Fraudulent_CreditCard_Transactions_Type/$
Import favorites
Amazon.co.uk - Onli...
Agoda
Express VPN
McAfee Security
LastPass password...
Gmail
YouTube
Maps

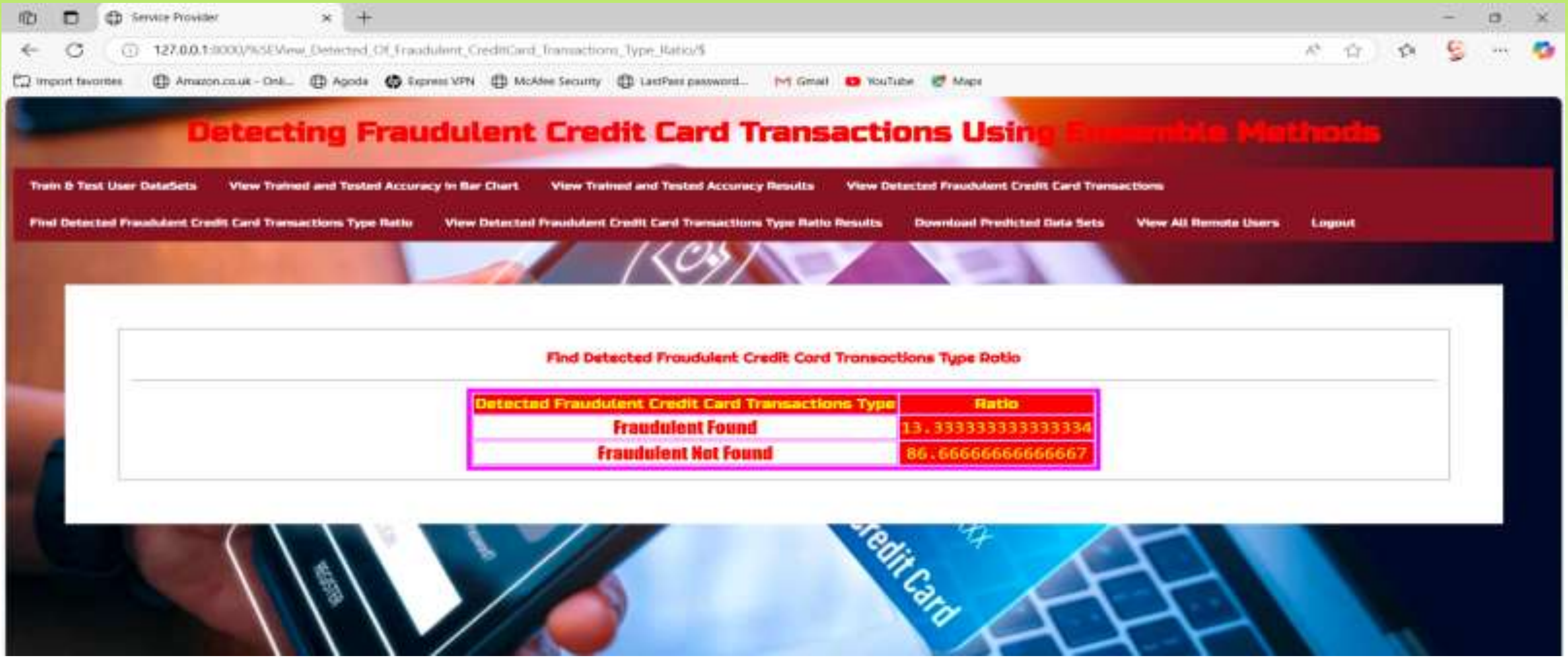
Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Train & Test User DataSets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Detected Fraudulent Credit Card Transactions
Find Detected Fraudulent Credit Card Transactions Type Ratio
View Detected Fraudulent Credit Card Transactions Type Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout

View Detected Froudulent Credit Cord Tronsoctions Type !!!

Fid	Trans_Date	CC_No	CC_type	Trans_Type	Amount	Firstname	Lastname	Gender
10.42.0.151-10.42.0.1-5732-53-17	21-04-22	4512830000000000000	First Wealth Credit Card	entertainment	2407.246035	Monica	Cohen	Female
175.100.207.231-10.42.0.211-80-37025-6	14-02-23	6011390000000000	Single Currency Card	home	311.963409	Natalie	Watson	Female
10.42.0.211-23.206.173.182-55689-80-6	16-01-22	4457730000000000000	EaseMyTrip	personal_care	100.302262	Stanley	Mayo	Male
184.86.36.20-10.42.0.151-443-33700-6	10-09-21	3600000000000000	EaseMyTrip	entertainment	140.983193	David	Everett	Male
5748394587438954389	02-03-2025	7859364829284893	MasterCard	Shopping	8000	Swathi	Pandre	Female

FRAUD DETECTION DETAILS

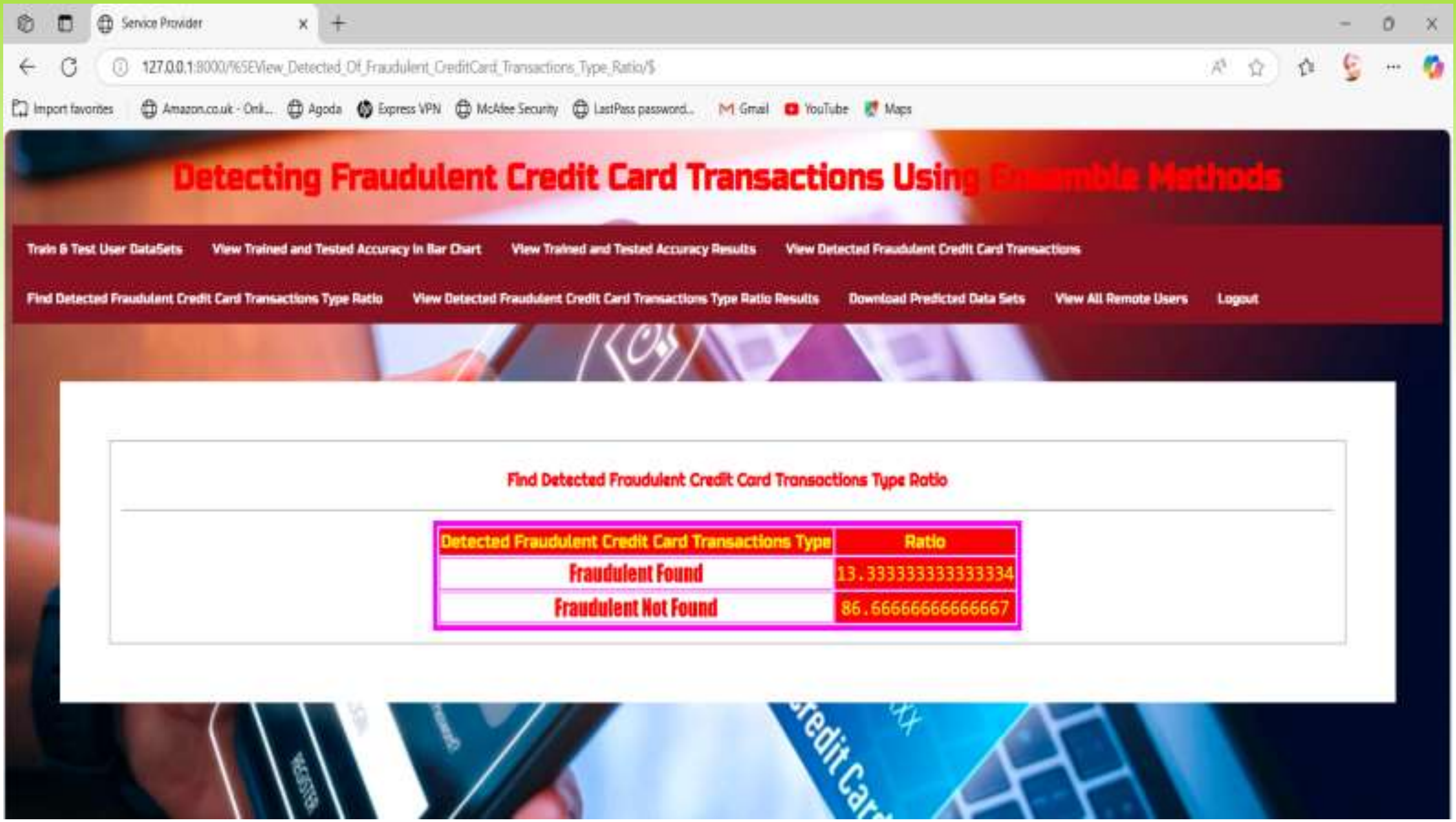

Service Provider
127.0.0.1:8000/%5EView_Detected_Of_Fraudulent_CreditCard_Transactions_Type_Ratio/$
Import favorites
Amazon.co.uk - Onl...
Agoda
Express VPN
McAfee Security
LastPass password...
Gmail
YouTube
Maps
Train & Test User DataSets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Detected Fraudulent Credit Card Transactions
Find Detected Fraudulent Credit Card Transactions Type Ratio
View Detected Fraudulent Credit Card Transactions Type Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout
Credit Card

Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Find Detected Fraudulent Credit Card Transactions Type Ratio

Detected Fraudulent Credit Card Transactions Type	Ratio
Fraudulent Found	13.333333333333334
Fraudulent Not Found	86.66666666666667

FRAUD DETECTION RATIO

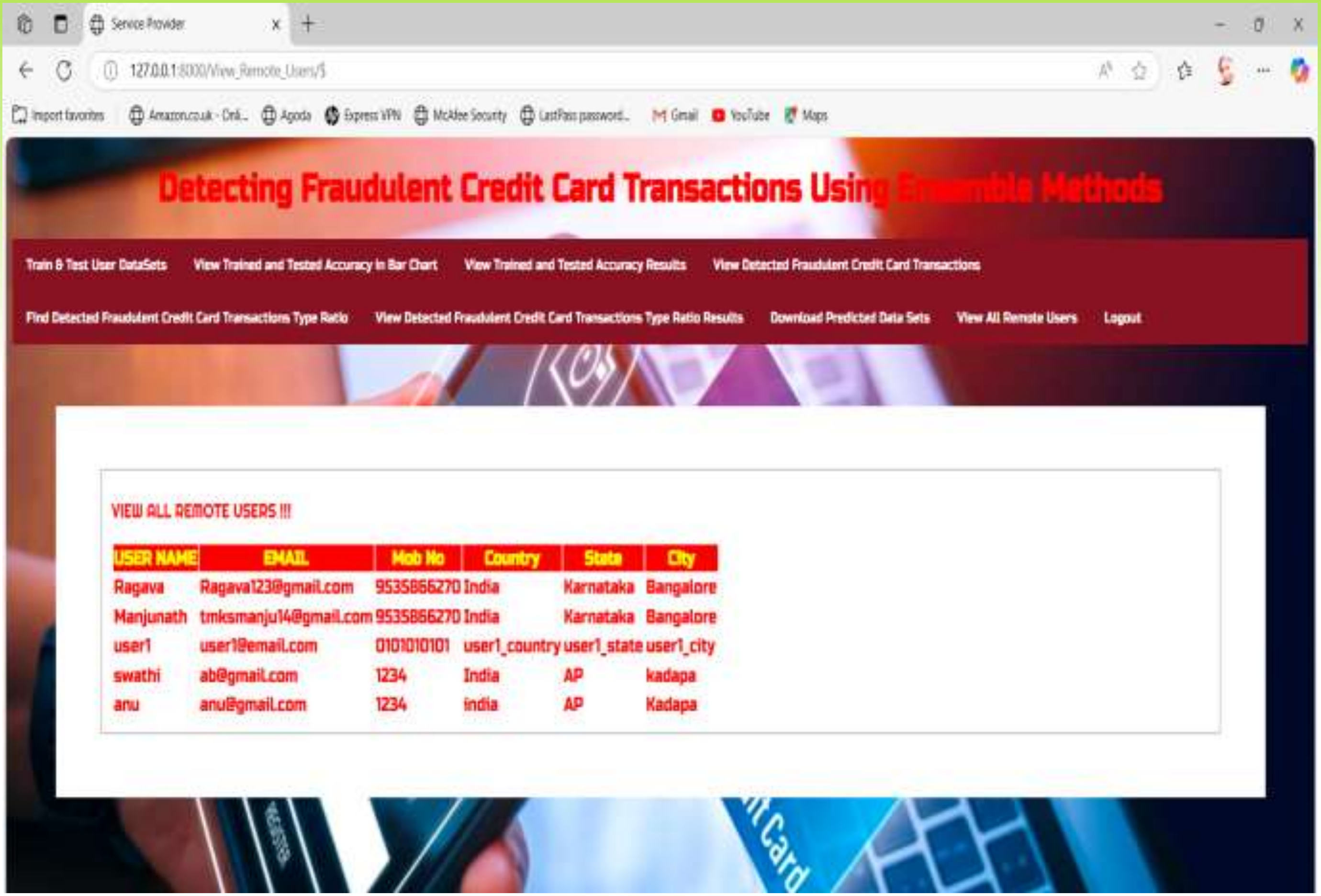

Service Provider
127.0.0.1:8000/%5EView_Detected_Of_Fraudulent_CreditCard_Transactions_Type_Ratio/$
Import favorites
Amazon.co.uk - Onl...
Agoda
Express VPN
McAfee Security
LastPass password...
Gmail
YouTube
Maps
Train & Test User DataSets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Detected Fraudulent Credit Card Transactions
Find Detected Fraudulent Credit Card Transactions Type Ratio
View Detected Fraudulent Credit Card Transactions Type Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout
Credit Card

Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Find Detected Fraudulent Credit Card Transactions Type Ratio

Detected Fraudulent Credit Card Transactions Type	Ratio
Fraudulent Found	13.333333333333334
Fraudulent Not Found	86.66666666666667

VIEW ALL REMOTE USERS



Service Provider
127.0.0.1:8000/View_Remote_Users/$
Import favorites
Amazon.co.uk - Onli...
Agoda
Express VPN
McAfee Security
LastPass password...
Gmail
YouTube
Maps

Detecting Fraudulent Credit Card Transactions Using Ensemble Methods

Train & Test User DataSets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Detected Fraudulent Credit Card Transactions
Find Detected Fraudulent Credit Card Transactions Type Ratio
View Detected Fraudulent Credit Card Transactions Type Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout

VIEW ALL REMOTE USERS !!!

USER NAME	EMAIL	Mob No	Country	State	City
Ragava	Ragava123@gmail.com	9535866270	India	Karnataka	Bangalore
Manjunath	tmksmanju14@gmail.com	9535866270	India	Karnataka	Bangalore
user1	user1@email.com	0101010101	user1_country	user1_state	user1_city
swathi	ab@gmail.com	1234	India	AP	kadapa
anu	anu@gmail.com	1234	india	AP	Kadapa

CONCLUSION

Ensemble methods like XGBoost and bagging outperform traditional classifiers in detecting fraudulent credit card transactions. However, they tend to overfit real datasets, reducing their generalizability. The strong performance on real data indicates that transaction approval follows a structured pattern that models can easily learn.

FUTURE WORK

Future research should focus on increasing data diversity and introducing randomness in authentication to prevent overfitting. Improving simulated data generation will help models generalize better across different fraud patterns. Enhancing model explainability using techniques like SHAP or LIME will provide better transparency in fraud detection.

REFERENCES

- Z. Faraji, “A review of machine learning applications for credit card fraud detection with a case study,” SEISENSE Journal of Management, vol. 5, no. 1, pp. 49–59, Feb. 2022.
- F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan, and M. Ahmed, “Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms,” IEEE Access, vol. 10, pp. 39700–39715, 2022.
- Nilson Report. Card Fraud Worldwide. Accessed: May 2023. [Online]. Available: https://nilsonreport.com/
- N. S. Alfaiz and S. M. Fati, “Enhanced credit card fraud detection model using machine learning,” Electronics, vol. 11, no. 4, p. 662, Feb. 2022.