**ONLINE PAYMENT FRAUD DETECTION**

AN INDUSTRY ORIENTED MINI REPORT

Submitted to

**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY,**

**HYDERABAD**

In partial fulfilment of the requirements for the award of the degree of

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

Submitted by

**PENCHALA HARIKA 20UK1A6631**

**SIRI CHANDANA GUNDA 20UK1A6654**

**LAXMAN VENNAPU 20UK1A6627**

**BANDI ARUNKUMAR 20UK1A6636**

Under the esteemed guidance of

**Mrs . R . SWATHI**

**( Assistant Professor )**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**VAAGDEVI ENGINEERING COLLEGE**

(Affiliated to JNTUH, Hyderabad)

Bollikunta, Warangal –

506005 **2020 – 2024**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINNERING**

**VAAGDEVI ENGINEERING COLLEGE**

**BOLLIKUNTA, WARANGAL – 506005**

**2023 – 2024**



**CERTIFICATE OF COMPLETION**

**INDUSTRY ORIENTED MINI PROJECT**

This is to certify that the UG Project Phase-1 entitled “ **ONLINE PAYMENT FRAUD DETECTION** ” Is being submitted by - ***PENCHALA HARIKA ( H.NO : 20UK1A6631 ), SIRI CHANDANA GUNDA ( H.NO : 20UK1A6654 ), LAXMAN VENNAPU ( H.NO : 20UK1A27 ), BANDI ARUN KUMAR ( H.NO : 20UK1A6636 )*** in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering to Jawaharlal Nehru Technological University Hyderabad during the academic year 2023-24, is a record of work carried out by them under the guidance and supervision.

**Project Guide Head of the Department**

**Mrs. R. Swathi**   **Dr. K. Sharmila**

(Assistant Professor) ( Professor )

**External**

**ACKNOWLEDGEMENT**

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to

our beloved **Dr.P.PRASAD RAO,** Principal, Vaagdevi Engineering College for making us

available all the required assistance and for his support and inspiration to carry out this UG

Project Phase-1 in the institute.

We extend our heartfelt thanks to **Dr.K.SHARMILA**, Head of the Department of CSM, Vaagdevi Engineering College for providing us necessary infrastructure and thereby giving us freedom to carry out the UG Project Phase-1.

We express heartfelt thanks to Smart Bridge Educational Services Private Limited, for their

constant supervision as well as for providing necessary information regarding the UG Project

Phase-1 and for their support in completing the UG Project Phase-1.

We express heartfelt thanks to the guide, **Mrs.R.SWATHI,** Assistant professor, Department of CSM for his constant support and giving necessary guidance for completion of this UG Project

Phase-1.

Finally, we express our sincere thanks and gratitude to my family members, friends for their

encouragement and outpouring their knowledge and experience throughout the thesis.

**PENCHALA HARIKA ( 20UK1A6631 )**

**SIRI CHANDANA GUNDA ( 20UK1A6654 )**

**LAXMAN VENNAPU ( 20UK1A6627 )**

**BANDI ARUNKUMAR ( 20UK1A6636 )**

**ABSTRACT**

The rapid digitization of financial transactions has led to an alarming rise in online payment fraud. Detecting and preventing these fraudulent activities have become paramount.

This abstract explores advanced techniques employed in online payment fraud detection. Leveraging machine learning algorithms, behavioural analytics, and real-time data integration, businesses and financial institutions can monitor transaction patterns, user behaviour , and device data instantaneously.

Predictive analytics and geolocation tracking enable the anticipation of potential fraud, allowing proactive measures. Furthermore, rule-based systems and biometric verification add layers of security, ensuring robust protection.

By integrating these sophisticated techniques, real-time fraud detection systems can swiftly identify anomalies and prevent unauthorized transactions, safeguarding the integrity of online payments and bolstering trust in digital financial systems.

**TABLE OF CONTENTS:-**

**1.INTRODUCTION ……………………………………………………….1**

* 1. **OVERVIEW…………………………………………………………....1**

* 1. **PURPOSE………………………………………………………………1**

**2.LITERATURE SURVEY………………………………………………..2**

**2.1 EXISTING PROBLEM………………………………………………..2**

**2.2 PROPOSED SOLUTION……………………………………………2-3**

**3.THEORITICAL ANALYSIS…………………………………………....4**

**3.1 BLOCK DIAGRAM……………………………………………………4**

**3.2 HARDWARE /SOFTWARE DESIGNING ………………………….4**

**4.EXPERIMENTAL INVESTIGATIONS ………………………………5**

**5.FLOWCHART……………………………………………………………6**

**6.RESULTS……………………………………………………………….7-9**

**7.ADVANTAGES AND DISADVANTAGES……………………………10**

**8.APPLICATIONS…………………………………………………………11**

**9.CONCLUSION…………………………………………………………...11**

**10.FUTURE SCOPE……………………………………………………….11**

**11.BIBILOGRAPHY ……………………………………………………...12**

**12.APPENDIX (SOURCE CODE) & CODE SNIPPETS…………….13-26**

**1.INTRODUCTION**

**1.1.OVERVIEW**

Online payment fraud detection is a vital component of digital security, aimed at identifying and preventing fraudulent activities in online financial transactions. As technology advances, so do the tactics of fraudsters.

To counter this, businesses and financial institutions employ sophisticated methods like machine learning algorithms, behavioural analytics, and real-time data analysis. These techniques scrutinize transaction patterns, user behaviour, and device data, allowing for the swift detection of anomalies and potential fraud.

By integrating predictive analytics, rule-based systems, biometric verification, and geolocation tracking, online payment fraud detection systems ensure the safety of digital transactions. This comprehensive approach not only protects businesses from financial losses and reputational damage but also safeguards consumers from identity theft and unauthorized transactions, bolstering trust in online payment systems.

**1.2.PURPOSE**

Online payment fraud detection serves as a critical shield in the digital realm, ensuring the legitimacy of financial transactions. Its purpose is to swiftly identify and thwart fraudulent activities using advanced technologies like machine learning and real-time monitoring. By doing so, it safeguards businesses from financial losses, maintains their reputation, and prevents legal complications.

Equally important, it protects consumers from identity theft and unauthorized charges, nurturing confidence in online transactions. This process is indispensable in upholding the trustworthiness of digital payment platforms, creating a secure environment for businesses and customers, essential for the modern digital economy.

**2. LITERATURE SURVEY**

**2.1 EXISTING PROBLEM**

Achieving efficiency in online payment fraud detection requires advanced machine learning algorithms, real-time data processing, and cross-channel integration. Enhancing accuracy to reduce false positives, coupled with continuous updates to tackle evolving fraud methods, can streamline detection. Collaboration among institutions and leveraging cutting-edge technologies ensure a more efficient and adaptive fraud detection system.

**2.2 PROPOSED SOLLUTION**

In response to the escalating challenges posed by online payment fraud, a multifaceted solution is proposed, combining advanced technologies, strategic collaborations, and dynamic process optimization. Central to this approach is the integration of cutting-edge machine learning algorithms. By implementing deep learning neural networks and anomaly detection models, the system gains the ability to analyse vast datasets in real-time. These algorithms, continuously trained with new data, possess the agility to discern intricate fraud patterns, enhancing accuracy and adaptability.

Real-time data processing lies at the core of this solution. Leveraging tools like Apache Kafka and Spark Streaming, the system can handle the rapid influx of transaction data, enabling instant analysis. Real-time processing is pivotal, ensuring that fraudulent activities are detected and addressed swiftly. Additionally, cross-channel integration is imperative. By aggregating data from various sources such as web transactions, mobile apps, and social media interactions, the system gains a comprehensive user profile.

* Advanced Machine Learning Algorithms
* Real-time Data Processing and Analysis
* Cross-Channel Integration and Behavioural Analysis
* Adaptive Rule-based Systems and Thresholds
* Collaborative Intelligence and Information Sharing

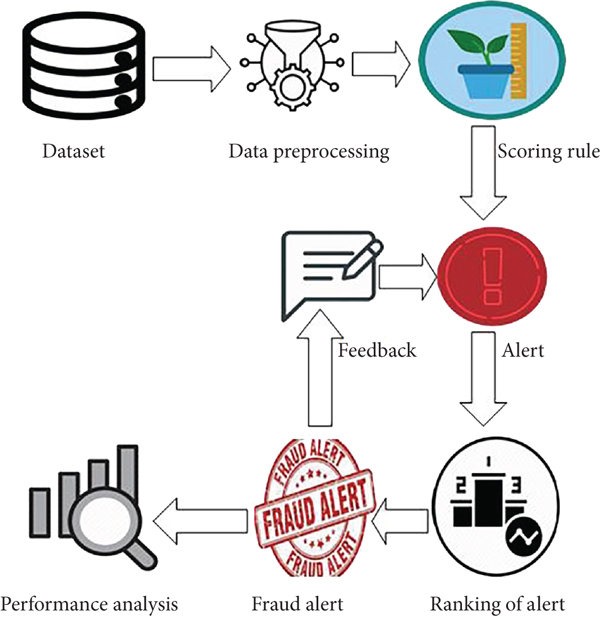
The proposed solution for online payment fraud detection integrates diverse strategies to ensure robust security measures in the digital realm. Cross-channel analysis provides a holistic view, identifying anomalies and inconsistencies indicative of fraudulent activities, especially when fraudsters exploit multiple platforms. Adaptive rule-based systems, combining predefined rules and machine learning insights, offer dynamic adaptability. These rules can be promptly updated in response to emerging fraud trends, significantly reducing false positives and enhancing overall accuracy.

Collaborative intelligence forms a crucial pillar, fostering strong partnerships among financial institutions and cybersecurity experts. This collaboration enables the swift sharing of threat intelligence, facilitating the rapid identification of emerging fraud schemes and the development of proactive countermeasures. Continuous process optimization plays a vital role, involving regular evaluations and feedback loops from real-world fraud incidents. This iterative approach ensures the system remains agile and effective, adapting to the ever-changing tactics employed by fraudsters.

In summary, this multifaceted solution combines cross-channel analysis, adaptive rule-based systems, collaborative intelligence, and continuous process optimization. By embracing these elements, businesses and financial institutions can establish an adaptive online payment fraud detection system. This proactive approach not only safeguards transactions by identifying anomalies and reducing false positives but also fosters industry-wide resilience, enhancing the digital financial ecosystem's security against evolving online payment fraud threats.

**3.THEORITICAL ANALYSIS**

**3.1 BLOCK DIAGRAM**



**3.2 HARDWARE / SOFTWARE DESIGNING**

**Hardware :**

* High-performance servers for data processing and analysis.
* Secure storage systems to store large datasets securely.

**Software :**

* **Data Analysis Tools** :

Python libraries like Pandas and NumPy for data manipulation.

* **Machine Learning Libraries** :

We have used Scikit-learn, seaborn, matplotlib.pyplot for implementing machine learning algorithms.

* **Real-time Processing** :

We have used train\_test\_split, sklearn.model\_selection for real-time data streaming.

* **Database Management** :

MySQL, or NoSQL databases for storing and retrieving transactional data efficiently.

* **Visualization Tools** :

Flask, pickle for visually representing patterns and trends

in the data.

Combining these elements, a robust Online Payment Fraud Detection system can be designed, capable of analysing vast amounts of data in real-time, detecting anomalies, and ensuring secure online transactions.

**4.EXPERIMENTAL INVESTIGATION**

In this project, we have used online payment fraud detection dataset. This dataset is a file consisting of labelled data and having the following columns:

1.**Type:** It contains different payment systems like Transfer, Debit, Cash\_out and Payments.

2.**Amount:** It contains the amount details.

3.**nameOrig:** This column typically represents the name or identifier of the originator (sender) of a transaction. This column contains information about the source of the funds or the party initiating the payment.

4.**oldbalanceOrg:** Thiscolumn in a dataset for online payment fraud detection typically represents the account balance before a transaction is initiated by the account holder. This column provides historical information about the account's balance prior to the transaction in question.

5.**newbalanceOrig:** This column provides important information about the account's balance following a transaction, and it can be useful for various purposes in the context of fraud detection.

6**.nameDest:** This column typically represents the recipient's account or destination account for a transaction.

7.**oldbalanceDest:** This column in a dataset for online payment fraud detection typically represents the recipient's account balance before a transaction is initiated.

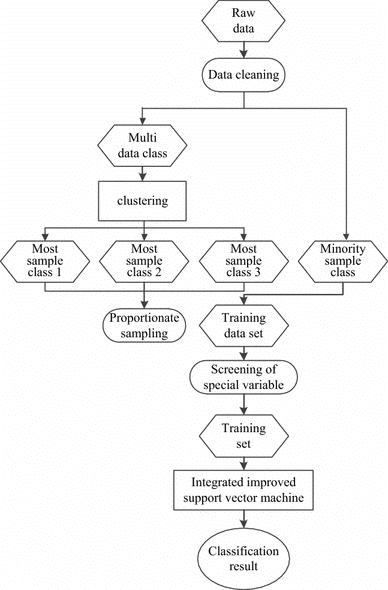
8.**newbalanceDest:** This column typically represents the recipient's account balance after a transaction is initiated and processed.

9.**isFraud:** This column in a dataset for online payment fraud detection is a binary column that typically indicates whether a transaction is fraudulent or not.

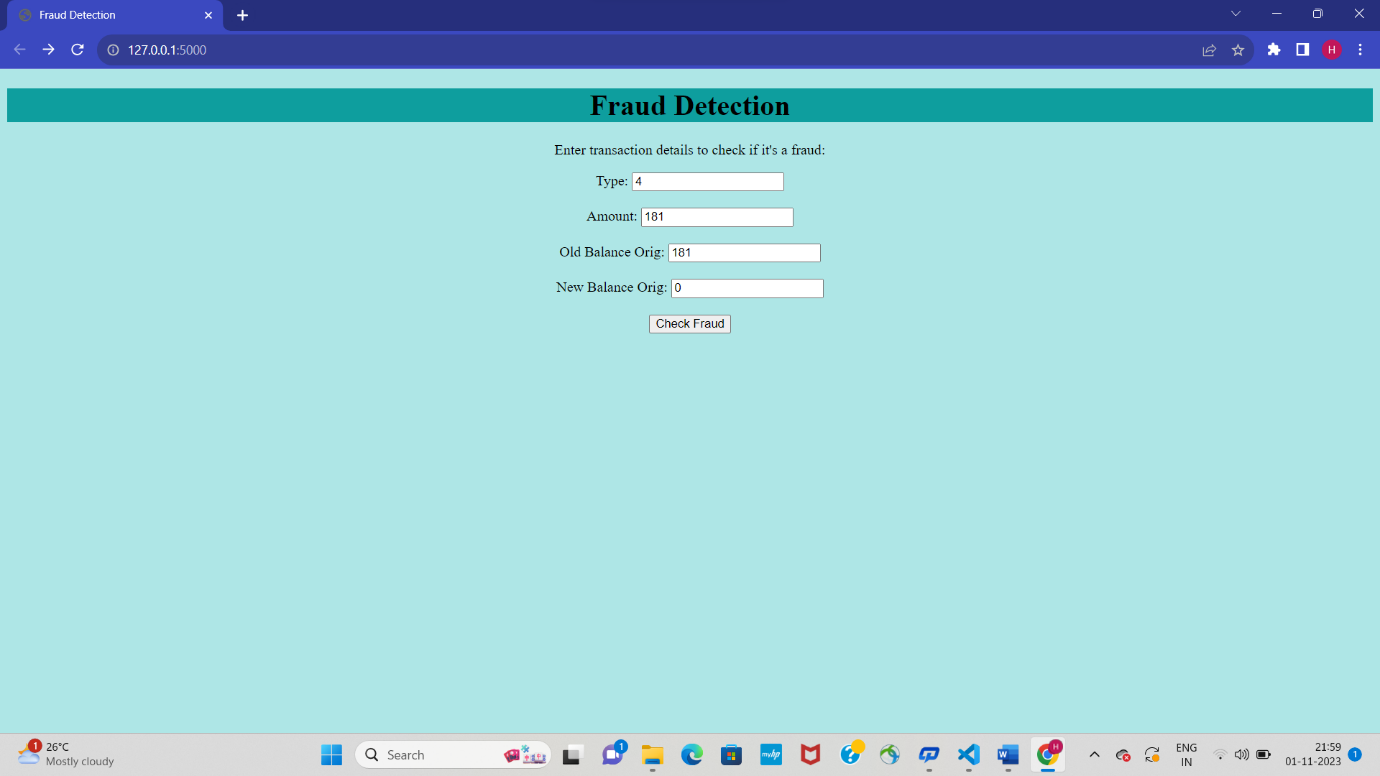
10.**isFlaggedFraud:** This column is used for specific cases where transactions exhibit highly unusual characteristics that may not be captured by standard fraud detection methods.

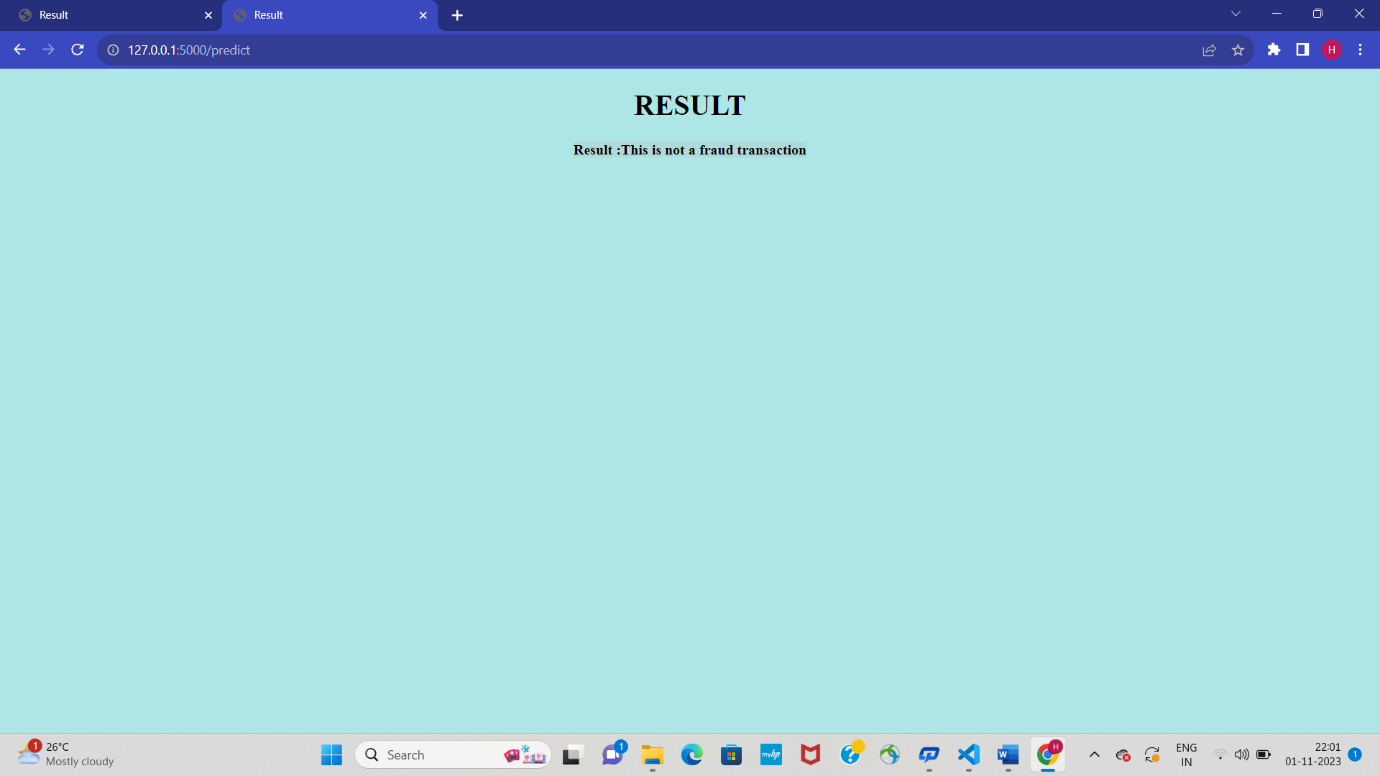
So we have used all this columns to detect whether the given data has a fraud transaction or not.

**5.FLOWCHART**

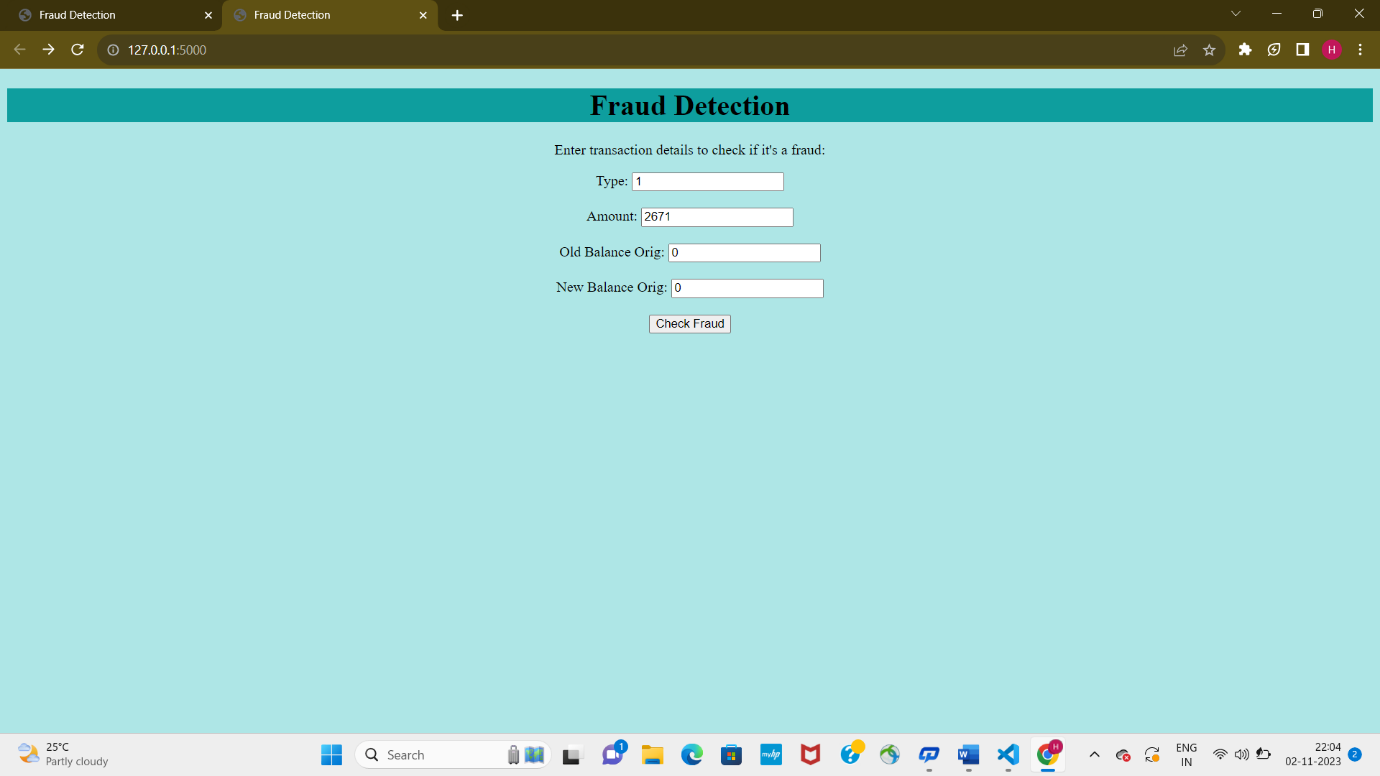


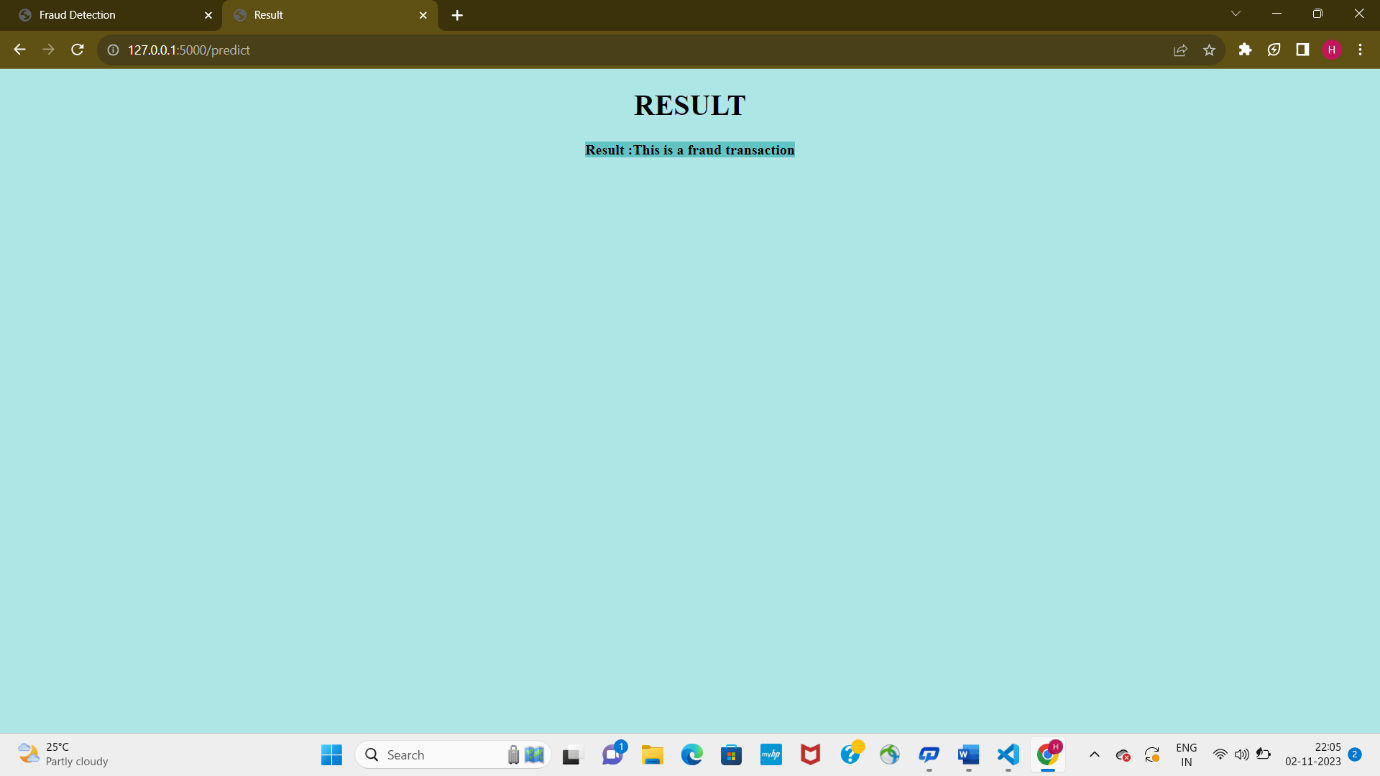
**6.RESULT**





This is not a Fraud Transaction





This is a Fraud Transaction

**7.ADVANTAGES AND DISADVANTAGES**

**ADVANTAGES**

* **Security:** Online payment fraud detection ensures the security of digital transactions, protecting businesses and consumers from financial losses.
* **Real-Time Detection**: Fraud detection systems operate in real-time, swiftly identifying and preventing fraudulent activities as transactions occur.
* **Reduced False Positives**: Advanced algorithms minimize false positives, distinguishing between genuine and fraudulent transactions more accurately.
* **Cross-Platform Protection**: These systems analyze transactions across various platforms, offering comprehensive security against fraudsters exploiting different channels.
* **Adaptability**: Fraud detection systems evolve to counter new fraud tactics, ensuring ongoing protection against emerging threats.
* **Customer Trust**: Establishes trust among consumers, encouraging online transactions and boosting confidence in digital payment methods.
* **Compliance:** Helps businesses adhere to industry regulations and standards, reducing legal liabilities related to financial transactions.
* **Cost Efficiency**: Prevents financial losses, reducing expenses associated with fraud-related incidents, legal fees, and reputation management.

**DISADVANTAGES**

* **Complex Implementation**: Setting up and maintaining sophisticated fraud detection systems can be complex, requiring significant investment in technology and skilled personnel.
* **Resource Intensiveness**: Continuous monitoring and analysis demand substantial computational resources, leading to high operational costs.
* **Technological Dependencies**: Dependence on advanced technologies makes systems vulnerable to technical glitches, potentially disrupting the detection process.
* **Privacy Concerns**: Collecting extensive user data for analysis raises privacy concerns, requiring stringent data protection measures to be in place.
* **Adaptation Challenges**: Fraudsters continually evolve tactics, challenging detection systems to keep up, necessitating constant updates and improvements.

**8. APPLICATIONS**

* **E-commerce Security**: Protects online stores from payment fraud, ensuring safe transactions for customers.
* **Banking Security**: Safeguards bank accounts by detecting unauthorized transactions and identity theft.
* **Mobile Payment Apps**: Secures transactions on smartphones and tablets, preventing fraud on digital wallets and payment apps.
* **Travel Booking Platforms**: Identifies and prevents fraudulent bookings on ticketing and travel websites.
* **Gaming Platforms**: Ensures fair play by detecting payment fraud in online gaming and gambling platforms.

**9. CONCLUSION**

In conclusion, online payment fraud detection is indispensable in the digital landscape, safeguarding financial transactions and ensuring the integrity of digital commerce. Through advanced technologies like machine learning and real-time data analysis, businesses and consumers are shielded from evolving fraud tactics. These systems not only prevent financial losses but also bolster customer trust, enabling seamless and secure online transactions.

However, it's essential for organizations to balance security with user experience, addressing false positives and ensuring swift, accurate detection. By embracing continual innovation, collaboration, and adherence to ethical standards, the digital realm can remain a secure space for financial interactions, fostering trust and reliability among users and businesses alike.

**10. FUTURE SCOPE**

The future of online payment fraud detection is promising, driven by advanced technologies. Artificial intelligence and machine learning will enhance predictive algorithms, identifying fraud patterns in real-time. Blockchain technology will ensure tamper-proof transaction records, bolstering security. Biometric authentication methods will add an extra layer of protection, ensuring secure user access. Additionally, the continuous evolution of these technologies, coupled with big data analytics, will lead to more accurate, adaptive, and efficient fraud detection systems, ensuring a safer digital payment environment for all stakeholders.

**11. BIBILOGRAPHY**

* Abdallah A, Maarof MA, Zainal A (2016) Fraud detection system: a survey. J Netw Comput Appl 68:90–113
* Ali A, Shukor AR, Siti HO, Abdu S (2022) Financial fraud detection based on machine learning: a systematic literature
* review. Review Appl Sci 12:9637
* Amiri M, Hekmat S (2021) Banking fraud: a customer-side overview of categories and frameworks of detection and
* prevention. J Appl Intell Syst Inf Sci 2(2):58–68
* Aggarwal CC, Sathe S (2017) Outlier ensembles: an introduction. Springer
* Bessis J (2011) Risk management in banking. Wiley, New York
* Bolton RJ, Hand DJ (2002) Statistical fraud detection: a review. Stat Sci 17(3):235–249
* Bolton RJ, Hand DJ (2001) Unsupervised proﬁling methods for fraud detection, Credit Scoring and Credit Control VII, pp 225.

**12.APPENDIX**

Model Building

* Dataset
* Google Collab
* Application Building
* HTML file
* CSV file

**SOURCE CODE**

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

import plotly.express as px

import pickle

from sklearn.model\_selection import train\_test\_split

from sklearn.svm import SVC

from sklearn.naive\_bayes import GaussianNB

from sklearn.tree import DecisionTreeClassifier

from sklearn.neighbors import KNeighborsClassifier

from sklearn.linear\_model import LogisticRegression

from sklearn import model\_selection

warnings.filterwarnings("ignore")

df=pd.read\_csv('onlinefraud.csv')

df.head(10)

df.columns

df.info()

df['step'].unique()

df.isnull().sum()

df.shape

df['type'].unique()

type=df['type'].value\_counts()

transaction=type.index

quantity=type.values

px.pie(df,values=quantity,names=transaction,hole=0.4,title="Distribution of transaction type")

df.replace(to\_replace=['PAYMENT', 'TRANSFER', 'CASH\_OUT', 'DEBIT', 'CASH\_IN'],value=[2,4,1,5,3],inplace=True)

type

df

df['isFraud']=df['isFraud'].map({0:'No Fraud',1:'Fraud'})

df.head(100)

x=df[['type','amount','oldbalanceOrg','newbalanceOrig']]

y=df.iloc[:,-2]

y

from sklearn.tree import DecisionTreeClassifier

from sklearn.linear\_model import LogisticRegression

xtrain,xtest,ytrain,ytest=train\_test\_split(x,y,test\_size=0.20,random\_state=42)

model.fit(xtrain,ytrain)

model.score(xtest,ytest)

model.score(xtest,ytest)\*100

model.predict([[2,9839,170136,160296]])

x

#models=[]

#results=[]

#names=[]

#from sklearn.linear\_model import LogisticRegression

#from sklearn.neighbors import KNeighborsClassifier

#from sklearn.naive\_bayes import GaussianNB

#from sklearn.svm import SVC

#models.append(('LR',LogisticRegression()))

#models.append(('KNN',KNeighborsClassifier()))

#models.append(('NB',GaussianNB()))

#models.append(('SVM',SVC()))

#models.append(('CART',DecisionTreeClassifier()))

#models #from sklearn import model\_selection

#for name,model in models:

# kfold=model\_selection.KFold(n\_splits=10,random\_state=42,shuffle=True)

#cv\_result=model\_selection.cross\_val\_score(model,x,y,cv=kfold,scoring='accuracy')

# results.append(cv\_result)

# names.append(name)

# meg='%s: %f'%(name,cv\_result.mean())

# print(meg)

import pickle

with open('fraud\_detection\_model.pkl', 'wb') as model\_file:

pickle.dump(model, model\_file, protocol=2)

**FLASK CODE**

from flask import Flask, request, render\_template

import pickle

app = Flask(\_name\_)

# Load the pickled model

with open('fraud\_detection\_model.pkl', 'rb') as model\_file:

model = pickle.load(model\_file)

@app.route('/')

def index():

return render\_template('index.html')

@app.route('/predict\_fraud', methods=['POST'])

def predict\_fraud():

if request.method =='POST':

type = int(request.form.get('type'))

amount = float(request.form.get('amount'))

oldbalanceOrg = float(request.form.get('oldbalanceOrg'))

newbalanceOrig = float(request.form.get('newbalanceOrig'))

prediction = model.predict([[type,amount,oldbalanceOrg,newbalanceOrig]])

# Return the prediction result

return render\_template('predict\_fraud.html', is\_fraud=prediction[0])

if \_name\_ == '\_main\_':

app.run(debug=True)

**CODE SNIPPETS**

**MODEL BUILDING**

