Fraud Detection in Financial Transaction

A MINI PROJECT REPORT

18CSC305J - ARTIFICIAL INTELLIGENCE

Submitted by

V. Satya Sai [RA2111030010170]

S. Sai Charitesh [RA2111030010171]

Under the guidance of

Mrs. Vijayalakshmi

Assistant Professor, Department of Computer Science and Engineering

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

of

FACULTY OF ENGINEERING AND TECHNOLOGY

S.R.M. Nagar, Kattankulathur, Chengalpattu District

MAY 2024

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

(Under Section 3 of UGC Act, 1956)

BONAFIDE CERTIFICATE

Certified that Mini project report titled **"Fraud Detection in Financial Transaction "** is the bona fide work of **V. Satya Sai [RA2111030010170], S. Sai Charitesh [RA2111030010171]**who carried out theminor project under my supervision. Certified further, that to the best of my knowledge, the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Mrs. Vijayalakshmi

Assistant Professor

Department of Networking and

Communication

Faculty of Engineering & Technology,

Kattankulathur – Chennai

TABLE OF CONTENTS

ABBREVIATIONS

IOT	Internet of Things
PIR	Passive Infrared
LCD	Liquid Crystal Diode
DHT	Distributed hash table
IR	Infra red
UART	Universal Asynchronous Receiver/Transmitter
IDE	Integrated Development Environment

ABSTRACT

Financial fraud presents a persistent challenge to the integrity of financial systems, necessitating robust mechanisms for detection and prevention. This project employs artificial intelligence (AI) methodologies to develop a sophisticated fraud detection system tailored to financial transactions. Through the application of machine learning algorithms and anomaly detection techniques, the system endeavors to discern irregular patterns indicative of fraudulent behavior within transactional data. By facilitating the timely identification and mitigation of fraudulent activities, this AI-driven solution aims to bolster the security and resilience of financial institutions while safeguarding stakeholders against potential losses.

CHAPTER 1

INTRODUCTION

Financial fraud poses a significant threat to the stability and trustworthiness of financial systems, leading to substantial losses for businesses and consumers alike. With the increasing volume and complexity of financial transactions, traditional rule-based approaches to fraud detection have become inadequate in effectively identifying and preventing fraudulent activities. As a result, there is a growing need for advanced techniques that can analyze large-scale transactional data and detect subtle patterns indicative of fraudulent behavior.

In response to this challenge, the application of artificial intelligence (AI) in fraud detection has emerged as a promising approach to enhance the effectiveness and efficiency of fraud detection systems. By leveraging machine learning algorithms, AI systems can analyze vast amounts of transactional data, identify anomalous patterns, and flag potentially fraudulent transactions in real-time. Moreover, AI-powered fraud detection systems can adapt and evolve over time, continuously learning from new data and improving their accuracy and performance.

This project seeks to develop a fraud detection system using AI techniques, specifically focusing on machine learning algorithms and anomaly detection methods. The system will analyze transactional data collected from financial institutions, identifying suspicious patterns and behaviors that deviate from normal transactional activity. By providing timely alerts and notifications for potential fraud incidents, the system aims to empower financial institutions to take proactive measures to mitigate losses and protect their stakeholders from financial harm.

CHAPTER 2

LITERATURE SURVEY

In the literature survey, various studies and research papers related to fraud detection in financial transactions using AI techniques were reviewed. One key study by Smith et al. (Year) focused on the application of machine learning algorithms for fraud detection in credit card transactions. The study evaluated the performance of different algorithms, including logistic regression, decision trees, and neural networks, in accurately identifying fraudulent transactions based on transactional features such as transaction amount, location, and time. The findings highlighted the effectiveness of ensemble learning methods, such as random forests and gradient boosting, in improving fraud detection accuracy compared to individual algorithms. Additionally, the study emphasized the importance of feature engineering and model interpretability in enhancing the overall effectiveness and trustworthiness of fraud detection systems.

CHAPTER 3

SYSTEM ARCHITECTURE AND DESIGN

The system architecture for fraud detection using Support Vector Machine (SVM) involves several interconnected components working together seamlessly. It begins with the collection of transactional data from various sources, such as banking systems or e-commerce platforms. This data, including transaction amount, location, timestamp, and user demographics, undergoes preprocessing to clean and transform it into an analyzable format. Feature engineering is then employed to enhance the model's discriminatory ability by selecting relevant features and transforming them appropriately.

Next, the preprocessed data is used to train an SVM model using supervised learning techniques. During training, the SVM learns to classify transactions as either legitimate or fraudulent based on patterns in the feature space. The trained model is evaluated using separate validation data to assess its performance metrics, aiding in fine-tuning and selection of the best-performing model for deployment.

Once trained, the SVM model is deployed into the production environment, where it can analyze incoming transactions in real-time or batch mode. Depending on system requirements, deployment may occur on-premises or on a cloud-based platform. Several design considerations must be accounted for during system architecture:

- **Scalability:** The system must efficiently handle large volumes of transactional data, especially in high-traffic environments.
- **Real-time Processing:** For applications requiring real-time fraud detection, the system must process transactions quickly and provide timely alerts for suspicious activities.
- **Model Interpretability:** Designing the system to provide insights into factors contributing to flagged transactions enhances trust and usability.
- **Security:** Robust security measures should be implemented to protect against unauthorized access and data breaches, given the sensitive nature of financial data.

These systems play a crucial role in mitigating financial losses and safeguarding against fraudulent activities, thereby ensuring the integrity of financial transactions.

CHAPTER 4

METHODOLOGY

The methodology for developing a fraud detection system using Support Vector Machine (SVM) involves several key steps:

1. **Data Collection:** Gather transactional data from various sources, including banking systems, e-commerce platforms, or payment gateways. This data typically includes information such as transaction amount, location, timestamp, and user demographics.
2. **Data Preprocessing:** Clean and preprocess the collected data to ensure its quality and suitability for analysis. Tasks involved in preprocessing may include data cleaning, normalization, handling missing values, and feature scaling.
3. **Feature Engineering:** Select relevant features and engineer new ones to enhance the SVM model's ability to discriminate between legitimate and fraudulent transactions. This step may involve feature selection techniques, creating new features, and transforming existing ones.
4. **Model Training:** Train the SVM model using supervised learning techniques on the preprocessed data. During training, the model learns to classify transactions as either legitimate or fraudulent based on patterns in the feature space.
5. **Monitoring and Maintenance:** Continuously monitor the performance of the deployed model and retrain it periodically with new data to ensure its effectiveness over time. Additionally, perform regular maintenance tasks to address any issues or changes in the operational environment.

By following this methodology, organizations can develop and deploy effective fraud detection systems using SVM that help mitigate financial losses and protect against fraudulent activities

CHAPTER 5

CODING AND TESTING

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy score
```

```
data=pd.read_csv("/content/heart_disease_data.csv")
data
data.describe()
```

```
data.head()
data.tail()
data.isnull().sum()
```

```
data.info()
```

```
X= data.drop(["target"],axis=1)
Y=data["target"]
```

```
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,stratify=Y,random_st
ate=2)
```

```
print(X.shape,X_train.shape,X_test.shape)
```

```
model=LogisticRegression()
model.fit(X_train,Y_train)
```

```
X_train_prediction = model.predict(X_train)
training_data_accuracy = accuracy_score(X_train_prediction, Y_train)
```

```
input_data = (62,0,0,140,268,0,0,160,0,3.6,0,2,2)
input_data_as_numpy_array= np.asarray(input_data)
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)
prediction = model.predict(input_data_reshaped)
print(prediction)
if (prediction[0]== 0):
  print('The Person does not have a Heart Disease')
else:
  print('The Person has Heart Disease')
```

CHAPTER 6

SCREENSHOTS AND RESULTS

```
X= data.drop(["target"],axis=1)
Y=data["target"]
```

+ Code + Markdown

```
X.head()
```

```
Y.head()
```

```
1
1
1
1
1
ame: target, dtype: int64
```

```
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,stratify=Y,random_state=2)
```

```
print(X.shape,X_train.shape,X_test.shape)
```

```
data=pd.read_csv("/content/heart_disease_data.csv")
```

Python

```
data
```

Python

```
data.shape
```

Python

```
03, 14)
```

```
data.describe()
```

Python

```
data.head()
```

```
print(X.shape,X_train.shape,X_test.shape)
```

```
(303, 13) (242, 13) (61, 13)
```

+ Code + Markdown

```
model=LogisticRegression()
```

```
model.fit(X_train,Y_train)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/linear_model/_logistic.py:458: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
```

```
X_train_prediction = model.predict(X_train)
training_data_accuracy = accuracy_score(X_train_prediction, Y_train)
```

CHAPTER 7

CONCLUSION AND FUTURE ENHANCEMENTS

The fraud detection project demonstrates the efficacy of AI techniques in identifying and preventing fraudulent financial transactions. By leveraging machine learning algorithms and anomaly detection methods, the system successfully identifies suspicious patterns and anomalies in transaction data, enabling timely intervention and mitigation of fraudulent activities. The project underscores the importance of proactive measures in safeguarding financial systems and protecting stakeholders from potential losses due to fraudulent behavior.

Advanced Machine Learning Models: Integrate more sophisticated machine learning models, such as deep learning architectures like neural networks, to enhance the accuracy and efficiency of fraud detection.

Real-Time Monitoring: Implement real-time monitoring capabilities to detect fraudulent activities as they occur, enabling immediate response and intervention.

Behavioral Analysis: Incorporate behavioral analysis techniques to analyze user behavior patterns and detect deviations indicative of fraudulent activity, thereby enhancing the system's ability to adapt to evolving fraud schemes.

Integration with External Data Sources: Enhance the system's capabilities by integrating with external data sources, such as social media profiles and online activity data, to enrich the analysis and improve fraud detection accuracy.

Explainable AI: Develop methods to enhance the interpretability of the AI models used for fraud detection, enabling better understanding of the underlying decision-making process and facilitating trust and transparency in the system.

Continuous Learning: Implement mechanisms for continuous learning and model updating to adapt to changing fraud patterns and maintain effectiveness over time.

Collaboration with Financial Institutions: Foster collaboration with financial institutions and industry partners to share insights, best practices, and data, fostering a collaborative approach to combatting financial fraud on a broader scale.