IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

FRAUD DETECTION IN BANKING DATA USING MACHINE LEARNING

1 ABDUL KABEER KHAN, 2 KATTA SRINIVAS, 3 MEKALA PAVAN KUMAR REDDY, 4 Mr. K. Jayaprasad , 5 Dr. Md Azizuddin

1,2,3 Student, 4 Principal / Guide

1,2,3 Department of Computer Science and Engineering

1,2,3,4 Mumtaz College of Engineering and Technology, Hyderabad, Telangana, India

[5]Department of Computer Science, Mumtaz College of Engineering and Technology, Hyderabad, Telangana, India

Abstract: The rapid growth of digital banking systems has led to an increase in fraudulent financial transactions, posing a serious threat to both customers and financial institutions. This paper presents a machine learning-based approach for detecting fraudulent transactions in banking data. The system utilizes various classification algorithms to analyze transaction patterns and identify anomalies that indicate potential fraud. The dataset is preprocessed to handle missing values, normalize features, and improve model performance. Techniques such as Logistic Regression, Decision Trees, and Random Forest are used to classify transactions as legitimate or fraudulent. The proposed model achieves high accuracy and precision, enabling early detection and prevention of fraud. The system helps financial institutions minimize financial losses and enhance security by providing a reliable and automated fraud detection mechanism.

Index Terms - Machine Learning, Fraud Detection, Banking Security, Classification Algorithms, Data Analysis

I. INTRODUCTION

In recent years, the banking sector has experienced significant growth due to digital transformation. However, this growth has also led to an increase in fraudulent activities, making fraud detection a critical concern. Traditional fraud detection systems are often rule-based and fail to detect complex patterns in large datasets.

Machine learning provides an efficient solution by analyzing transaction behavior and identifying anomalies in real-time. This project focuses on developing a fraud detection system that uses machine learning algorithms to improve detection accuracy and reduce financial losses. The system is designed to enhance security and ensure safe digital transactions.

II. Related Work

Various fraud detection techniques have been proposed using data mining and machine learning approaches. Traditional systems rely on predefined rules, which are ineffective against evolving fraud patterns. Recent research focuses on supervised and unsupervised learning techniques for anomaly detection.

Machine learning models such as Decision Trees, Random Forest, and Neural Networks have shown promising results in detecting fraudulent transactions. These models analyze transaction behavior and identify deviations from normal patterns. However, challenges such as imbalanced datasets and false positives still exist. This project aims to improve detection accuracy using optimized machine learning techniques.

III. System Design

The proposed system is designed to detect fraudulent transactions using machine learning techniques. It follows a structured workflow consisting of data collection, preprocessing, model training, and prediction.

System Architecture:

• Data Collection: Transaction data is collected from banking datasets.

• Data Preprocessing: Cleaning, normalization, and feature selection are performed.

• Model Training: Machine learning algorithms such as Random Forest and Logistic Regression are trained.

• Prediction Module: The system classifies transactions as fraudulent or legitimate.

Modules:

• Data Preprocessing

• Feature Extraction

• Model Training

• Fraud Detection

• Result Visualization

Key Functional Modules

The LMS consists of multiple modules working together:

- Course Management
- Assignment & Quiz Management
- Attendance Tracking
- AI Recommendation Engine
- Chatbot System
- Analytics Dashboard

Table 1 Core modules of Website

Module	Function
Data Collection	Collects transaction data
Preprocessing	Cleans and prepares data
Feature Extraction	Identifies important features
Model Training	Trains ML algorithms
Fraud Detection	Detects fraudulent transactions
Result Analysis	Displays results

IV. Implementation Details

The system is implemented using Python and machine learning libraries such as Scikit-learn, Pandas, and NumPy. Data preprocessing techniques are applied to handle missing values and normalize the dataset. The model is trained using classification algorithms, and performance is evaluated using accuracy, precision, and recall.

The system provides a user-friendly interface for analyzing transaction data and detecting fraud efficiently.

V. Preliminary Evaluation And Discussion

The system was tested on banking transaction datasets and achieved high accuracy in detecting fraudulent transactions. The Random Forest model performed better compared to other algorithms.

The system effectively reduces false positives and improves detection efficiency. However, performance depends on dataset quality and size.

VI. Conclusion

The proposed fraud detection system successfully demonstrates the use of machine learning in identifying fraudulent transactions. It improves security and reduces financial losses. Future work can include deep learning techniques and real-time fraud detection systems.

VII. Acknowledgment

The authors express their sincere gratitude to the faculty and management of Mumtaz College of Engineering and Technology. Special thanks to the project guide for valuable support and guidance.

References

[1] Scikit-learn Documentation – https://scikit-learn.org
[2] Kaggle Dataset – https://kaggle.com
[3] Python Documentation – https://python.org
[4] IEEE Papers on Fraud Detection