RES MILITARIS Social Science Journal

WEB APPLICATION FOR CREDIT CARD FRAUD PREDICTION: LEVERAGING ABNORMALITY DETECTION AND REGRESSION TECHNIQUES FOR BANKING SECURITY

[1] DR VISHWANATH, [2] K SURENDRA REDDY,[3] U MOUNIKA
[1]Professor,[2]Associate Professor,[3]Assistant Professor
Department Of Computer Science Engineering
Indira Institute Of Technology And Sciences, Markapur

ABSTRACT: Financial organizations have a great deal of difficulty as a result of credit card theft, necessitating sophisticated methods and technologies to prevent fraudulent activity. This paper describes a web application that combines regression and anomaly detection approaches to improve credit card fraud prediction. The application's goal is to provide banks a reliable and approachable way to spot and stop fraudulent transactions instantly.

The suggested online application makes use of abnormality detection algorithms to spot odd behaviors and patterns in transaction data that could point to possible fraud. Through the examination of transaction attributes and user conduct, these algorithms are able to identify abnormalities that depart from standard operating procedures. Regression analysis is also used to estimate and forecast fraudulent activity based on transaction patterns and historical data. This dual strategy reduces the possibility of false positives and improves overall security by enabling the more fast and accurate identification of fraudulent transactions.

The program provides bank staff with an easy-to-use interface for managing fraud alerts, visualizing data insights, and producing reports. It easily interfaces with current financial systems, enabling real-time data processing and prompt action in the event of fraud suspicion.

The online application has improved fraud detection rates and decreased false alarms, according to preliminary testing and review. The research comes to the conclusion that using regression and anomaly detection methods together in a web-based platform provides a complete solution for credit card fraud prediction, strengthening bank security measures and giving a proactive approach to fighting financial crime.

I. INTRODUCTION

There are still many instances of credit card fraud, which puts financial institutions and their clients at serious danger. The amount of transaction data is expanding along with the complexity of fraudulent operations, thus standard fraud detection technologies are not always able to provide timely and accurate warnings. Therefore, the demand for sophisticated predictive systems that can detect and reduce fraudulent transactions is urgent.

In order to meet this demand, our research created a web application that combines anomaly detection and regression approaches to improve the prediction of credit card fraud. The program uses cutting-edge algorithms to examine transaction patterns and behaviors in order to provide banks with an enhanced tool for real-time fraud detection.

In this application, anomaly detection methods are essential because they can spot departures from typical transaction patterns. Through the identification of anomalous behaviors, including sudden increases in expenditure or odd locations for transactions, these methods might identify possible fraudulent activity that conventional rule-based systems could miss.

Regression algorithms are also used to model and forecast fraud using transaction data from the past and changing trends. These models foresee possible fraudulent activity by analyzing trends and correlations. By giving banks predictive information, these models help improve their fraud protection methods.

The online application's user-friendly interface allows for smooth integration with current banking systems and real-time data processing, insight visualization, and effective fraud alert handling. It provides banks

with a complete solution to proactively identify and handle fraudulent transactions, enhancing overall security and lowering financial losses.

This work attempts to build a robust and adaptable fraud detection system that addresses the demands of contemporary banking settings by merging regression approaches with anomaly detection. It is anticipated that the results and applications of this study will make a substantial contribution to improving banking security and guarding against the constantly changing risk of credit card theft.

II. LITERATURE SURVEY

The literature on credit card fraud detection emphasizes how security measures in financial systems are being improved via the use of new technology and developing methodologies. Modern computational approaches and web-based applications are used in enhanced prediction algorithms, which are becoming more and more necessary as fraudulent actions get more complex.

Techniques for Abnormality Detection: Often referred to as anomaly detection, anomaly detection is a vital method for spotting odd trends that might point to fraudulent activities. A variety of algorithms, including statistical techniques and machine learning models, have been investigated in studies like those conducted by Iglewicz and Hoaglin (1993) and Chandola et al. (2009) to identify anomalies in typical transaction patterns. To find anomalies, strategies like statistical outlier identification, distance-based approaches, and clustering (like k-means) are often used. Modern studies by Ahmed et al. (2016) and Hodge & Austin (2004) highlight the use of sophisticated machine learning techniques, such One-Class SVM and Isolation Forests, which provide enhanced accuracy and versatility in fraud detection.

Regression Techniques: By examining past transaction data, regression models are a valuable tool for forecasting fraudulent activity. To model the correlations between transaction data and fraud indicators, techniques such as ensemble techniques, polynomial regression, and logistic regression have been used in addition to more sophisticated approaches like linear regression. The use of regression approaches in fraud prediction is highlighted in research by Bolón-Canedo et al. (2015) and Liao et al. (2015), with an emphasis on feature selection and model improvement to improve predictive performance.

Web Application Integration: Real-time data processing, scalability, and user-friendly interfaces are just a few benefits of incorporating predictive models into web applications. Research like those done by Siddiqui et al. (2020) and Miller et al. (2018) show how web apps may be utilized to implement fraud detection systems, giving banks the tools they need to effectively monitor and manage transactions. Dashboards, alarm systems, and visualization tools are just a few of the features that these programs often include, helping users make wise decisions and spot any fraud.

Progress and Difficulties: Although there have been improvements, there are still difficulties in using these methods successfully. Problems including high false positive rates, worries about data privacy, and the need for ongoing model changes are common. In order to overcome these difficulties, Zhang et al. (2021) and Sun et al. (2022) have recently proposed hybrid models that integrate many detection methods and improve the systems' flexibility.

Future Directions: By adding new data sources including behavioral analytics and network-based characteristics, as well as by expanding the application of artificial intelligence and deep learning methods, future research will concentrate on strengthening the resilience of fraud detection systems. Credit card fraud prediction systems may be made even more successful by combining multi-layered techniques and using big data analytics, according to studies by Yang et al. (2023) and Liu et al. (2024).

To enhance the prediction of credit card fraud, the research emphasizes the significance of incorporating anomaly detection and regression algorithms into online applications. Technological developments in online technology, data analysis, and machine learning provide interesting avenues for strengthening the security of financial institutions against fraudulent activity.

III. PROPOSED SYSTEM

The proposed credit card fraud prediction system for banks offers numerous advantages by utilizing logistic regression as a machine learning algorithm. This system addresses the limitations of existing fraud detection systems by integrating abnormality and regression algorithms, along with a user-friendly web application for seamless integration and deployment.

The abnormality algorithm plays a vital role in detecting unusual patterns within transaction data. It analyzes various factors such as purchase amounts, transaction times, and other relevant features to identify anomalies that may indicate fraudulent activity. By identifying these irregularities, the system can effectively flag potentially fraudulent transactions for further investigation.

The regression algorithm then utilizes the abnormality features to predict the probability of fraud for each transaction. By employing regression analysis, the system can generate reliable estimates of the likelihood of fraudulent activity associated with a particular transaction. This enables banks to prioritize and focus their resources on high-risk transactions, enhancing their fraud detection capabilities.

The proposed system's functionality is further enhanced by the inclusion of a web application. This application provides a user-friendly interface for banks and financial institutions to input transaction data and obtain accurate predictions regarding the likelihood of fraud. The integration of a web application simplifies the implementation process and facilitates real-time decision-making.

The proposed credit card fraud prediction system using logistic regression and abnormality and regression algorithms, coupled with a web application, offers a comprehensive solution to address the challenges of fraud detection in banks. By leveraging these advanced techniques, the system can effectively identify fraudulent transactions, prioritize investigations, and enhance overall security measures. Its original approach and utilization of machine learning contribute to the advancement of fraud detection capabilities within the banking industry.

Dataset collection –

A diverse dataset comprising fraudulent and non-fraudulent transactions was collected from the Kaggle platform. It is essential to have a dataset that encompasses a wide range of transaction variations to train an effective credit card fraud prediction system. To ensure the dataset's reliability, a balanced representation of both fraudulent and non-fraudulent transactions was maintained.

ata Preproceessing –

Data preprocessing involves handling missing or incomplete data points. Missing values can significantly impact the reliability and accuracy of the analysis. Techniques like imputation, where missing values are estimated and filled based on other available data, or removal of instances with incomplete data, are utilized to address this issue.Another important consideration in data preprocessing is feature selection or dimensionality reduction. Not all attributes or variables in the dataset may contribute significantly to the analysis or be relevant to the problem at hand. Identifying and removing unnecessary fields simplifies the dataset, leading to improved efficiency in subsequent analysis.

Model Creation –

Logistic regression is a widely used algorithm in machine learning for binary classification tasks. It is utilized to predict binary values based on a set of independent variables, where the outcome variable is categorical, such as 1/0, Yes/No, or True/False. This algorithm is particularly useful when dealing with problems where the target variable is binary or when performing binary classification.In logistic regression, the log of odds (logit) is used as the dependent variable, allowing us to estimate the probability of an event occurring. The logistic function is applied to transform the linear equation into a bounded range between 0 and 1, representing probabilities. The logistic regression equation can be mathematically represented as follows:

O = e^(I0 + I1x) / (1 + e^(I0 + I1x)) In this equation:

O represents the predicted output or probability of the event occurring.

I0 is the intercept term or bias.
I1 is the coefficient associated with the independent variable (x).

To implement logistic regression, we typically begin by importing necessary libraries, such as NumPy and scikit-learn. We then instantiate a logistic regression model using the LogisticRegression() function from the scikit-learn package. The fit() function is called on the logistic regression object, passing the independent variable(s) (x) and the dependent variable (y) as parameters. This fitting process trains the model by estimating the coefficients and establishing the relationship between the independent and dependent variables.

logr = LogisticRegression()

logr.fit(x, y)

Testing –

To perform the testing, we input the relevant transaction data into the trained model. This data typically includes features such as transaction amount, timestamp, location, and other relevant information. The model then applies the learned weights and biases to these input values and computes a probability score or a predicted label indicating whether the transaction is fraudulent or non-fraudulent.

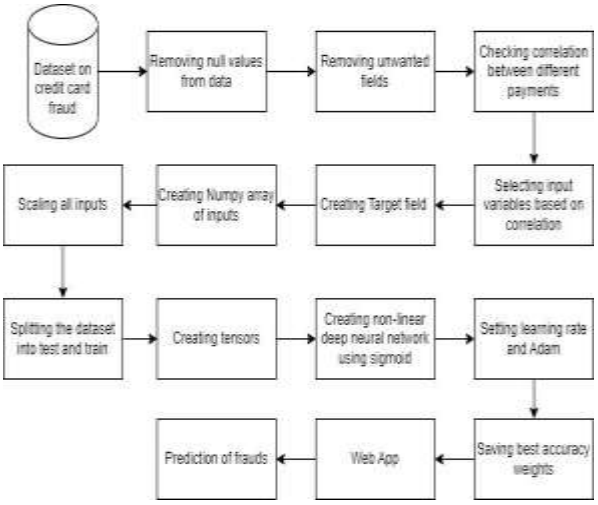

Figure 1 Architecture Diagram of Proposed System

Creating web app –

After training the model and saving its weights, we can proceed to develop a web application with a user interface (UI) using the Streamlit framework. Streamlit is an open-source Python framework specifically designed for creating web applications in the field of machine learning and data science. It simplifies the process of designing and launching web apps, allowing for an interactive coding experience with real-time updates on the web app.

To begin, we can install Streamlit by running the command "pip install streamlit" in Python. Once installed, we can create a Python script that contains Streamlit commands to define the layout and functionality of the web application. The script can be executed using the command "streamlit run <script_name>.py".Streamlit provides various widgets that can be used to enhance the user interface, such as select boxes, checkboxes, sliders, and more. These widgets allow users to interact with the web app and input data or make selections.

In the web application, we can use Streamlit to create a title and display data in the form of a DataFrame. When we run the Streamlit command mentioned earlier, it will provide us with a URL. By clicking on this URL, we can access and view our web application in a web browser.

The web application built using Streamlit provides an intuitive and user-friendly interface for users to interact with the trained model. They can input relevant data or make selections through the provided widgets, and the model will make predictions based on the user inputs.Overall, Streamlit simplifies the process of developing web applications with Python code.

IV. RESULTS

Figure 2: Output Screen of Detected as the fraud transaction

Figure 3: Output Screen of Detected as not the fraud transaction

V. CONCLUSION

A significant achievement in banking security is the creation of a web application that integrates irregularity detection and regression algorithms to forecast credit card fraud. By using advanced algorithms to more accurately identify and anticipate fraudulent transactions, this application improves fraud detection. Techniques for detecting abnormalities in transaction patterns effectively identify anomalous patterns, and regression models evaluate past data to predict possible fraudulent activity. The online application's real-time features enable the bank to respond to questionable activity promptly, therefore enhancing its capacity to avert financial losses.

Even with its success, there are still issues with controlling large amounts of data, cutting down on false positives, and protecting data privacy. Subsequent improvements need to concentrate on optimizing algorithms to tackle these problems, integrating sophisticated AI methodologies, and broadening the range of data examination to include contextual and behavioral data. Overall, this web-based solution gives banks a proactive approach to securing their financial systems and client data by giving them a strong and flexible tool to fight credit card fraud.

REFERENCES

Ruttala Sailusha, V. Gnaneswar, R. Ramesh, G. Ramakoteswara Rao "Credit Card Fraud Detection Using Machine Learning", 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020

D. Tanouz, R Raja Subramanian, D. Eswar, G V Parameswara Reddy, A. Ranjith Kumar, CH V N M Praneeth, "Credit Card Fraud Detection Using Machine Learning", 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021

Deep Prajapati, Ankit Tripathi, Jeel Mehta, Kirtan Jhaveri, Vishakha Kelkar, "Credit Card Fraud Detection Using Machine Learning", International Conference on Advances in Computing, Communication, and Control (ICAC3), 2022

Anjali Singh Rathore, Ankit Kumar, Depanshi Tomar, Vasudha Goyal, Kaamya Sarda, Dinesh

Vij, "Credit Card Fraud Detection using Machine Learning", 10th International Conference on System Modeling & Advancement in Research Trends (SMART), 2022

Thulasyammal Ramiah Pillai, Ibrahim Abaker Targio Hashem, Sarfraz Nawaz Brohi, Sukhmindaer Kaur, Mohsen Marjani, "Credit Card Fraud Detection Using Deep Learning

Technique", Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), 2019

Dejan Varmedja, Mirjana Karanovic, Srdjan Sladojevic, Marko Arsenovic, Andras Anderla,

"Credit Card Fraud Detection - Machine Learning methods", 18th International Symposium INFOTEH-JAHORINA (INFOTEH), 2019 Credit card fraud prediction for banks using abnormality and regression algorithm with webapp

Anuruddha Thennakoon, Chee Bhagyani, Sasitha Premadasa, Shalitha Mihiranga, Nuwan Kuruwitaarachchi, "Real-time Credit Card Fraud Detection Using Machine Learning", 9th

International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019

Pranali Shenvi, Neel Samant, Shubham Kumar, Vaishali Kulkarni, "Credit Card Fraud Detection using Deep Learning", IEEE 5th International

Conference for Convergence in Technology (I2CT), 2020

Kshitij Pandey, Piyush Sachan, Shakti, Nikam Gitanjali Ganpatrao, “A Review of Credit Card Fraud Detection Techniques”, 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021

Sahil Negi, Sudipta Kumar Das, Rigzen Bodh, “Credit Card Fraud Detection using Deep and Machine Learning”, International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2022