IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, May 2026
Impact Factor: 8.2

AI-Based Credit Card Anomaly Detection Using Z-Score and Isolation Forest

Manoj Shinde[1], Preksha Pokharna[2], Sahil Ramteke[3], Prashakha Mishra[4], Shivraj Patil[5]
Assistant Professor, Department of Computer Science[1]
Students, Department of Computer Science[2,3,4,5]
MIT ADT University, Pune, India

Abstract: *With the increasing flexibility of online money flow, there is increased potential of fraud occurring in such money transfers, especially using credit cards. It has therefore become essential to detect financial fraud, since the prevention of financial loss for the bank depends on it. This paper provides an overview of a method for detecting financial anomaly detection using machine learning. The method consists of two statistical techniques and machine learning techniques that will be used to detect financial fraud, with the two main techniques being the Z-Score and Isolation Forest technique. The Z-Score measures the deviation of transaction values from their normal range, whereas the Isolation Forest detects outliers by isolating them from the rest of the dataset. The source of data collection will consist of public credit card transaction datasets. Python will be used for implementing the machine learning process.*

Keywords: Anomaly Detection, Machine Learning, Credit Card Fraud Detection, Isolation Forest, Z-Score

I. INTRODUCTION

As a result of development of digital payment systems, the use of credit cards became much more widespread. Although such progress has greatly improved the experience of making payments for people, it also brought some negative aspects, such as the rising threat of fraud. Credit card fraud represents one of the most common financial crimes, resulting in significant losses to many banks, firms, and private individuals. However, current methods of fraud detection based on rules or manual monitoring often prove ineffective and unable to detect novel instances of fraud.

The use of machine learning can be an efficient method for identifying credit card fraud due to its ability to detect unusual patterns in transaction data autonomously. Anomaly detection refers to a method used to detect unusual or rare patterns in datasets, which may indicate various types of anomalies. Anomalies are any observations that differ significantly from the norm.

The purpose of this research is to develop an anomaly detection system utilizing machine learning models to identify any suspicious transactions made using credit cards. In order to do so, this system will make use of Z-scores to detect statistical outliers and Isolation Forest technique to identify anomalies by partitioning the data. This process allows the detection of unusual transaction without access to fraud data labels.

The main aim of this research is to develop a system for analyzing transactions data and identifying abnormal patterns.

II. TECHNICAL APPROACH

The proposed system uses machine learning techniques to detect anomalous credit card transactions. The methodology involves several steps including data collection, preprocessing, anomaly detection, and analysis of results.

DOI: 10.48175/IJARSCT-35157

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, May 2026
Impact Factor: 8.2

A. Dataset

The dataset used in this project is from a publicly accessible credit card transaction database. This database has many transactions along with their respective features which include transaction amount, transaction time, and other features that are anonymized to help detect anomalies.

B. Data Preprocessing

Before applying machine learning techniques on the dataset, we need to preprocess the dataset. The following tasks are performed during data preprocessing:

- Dealing with missing values
- Normalizing data
- Selecting features
  - Removing irrelevant attributes

C. Z-Score Analysis

Z-Score refers to a statistical technique that tells us about how far away a data point is from the average of a dataset. In simple terms, the Z-score tells us how many standard deviations away the data point is from its average value.

If a transaction has a very high or very low Z-Score value, it means that it is highly unlikely from a normal transaction. The Z-Score allows us to identify transactions that have extremely high or low values compared to usual transactions.

D. Isolation Forest Algorithm

The Isolation Forest algorithm refers to an unsupervised machine learning algorithm that is used for anomaly detection. It randomly selects features and splits the data point into subsets. As the anomalies are few and unique to regular data points, they get separated quickly.

Every data point has an anomaly score that identifies anomalies in the dataset. Transactions that have high anomaly scores are considered suspicious transactions.

E. System Workflow

The following is the general workflow of the system that we use to detect transaction anomalies. This workflow includes the following steps:

- Loading credit card transactions database
- Data preprocessing and data cleaning
- Applying Z-Score analysis to detect statistical anomalies
- Training isolation forest model with transaction data
- Generating anomaly scores of transactions
- Identifying suspicious transactions

III. RESULTS AND DISCUSSION

Anomaly Detection System was developed using Python language employing different machine learning packages like Pandas, NumPy, and Scikit-learn. It was tested on a dataset including several credit card transactions.

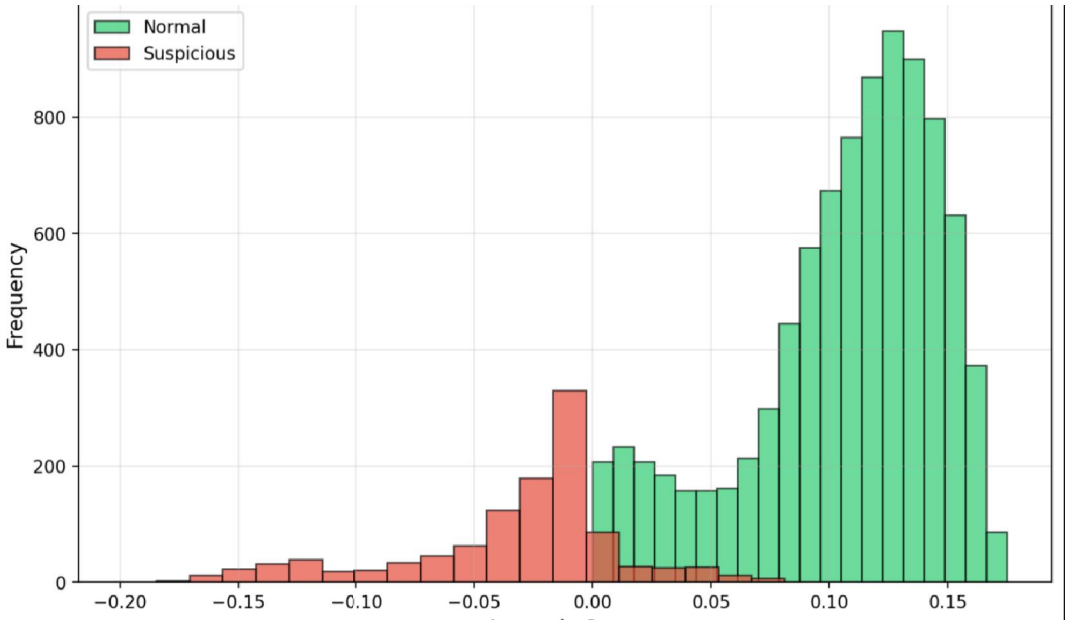
For each individual transaction, an anomaly score is determined based on how far the data point deviates from the norm. Transactions having a high anomaly score are considered possible outliers.

According to results, the Isolation Forest technique is successful in recognizing unusual transaction patterns. Z-Score technique proves helpful in identifying outliers in the dataset which might represent unusual activity.

The anomaly detection system may be useful for financial organizations to detect abnormal transactions. It can make the process easier for financial institutions as they do not have to spend much time manually analyzing their datasets.

Graphs can also be plotted for visualizing the data and observing normal and abnormal data distributions in a dataset.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, May 2026
Impact Factor: 8.2



IV. CONCLUSION

In this study, a machine learning-based system that detects anomalies in credit card transactions has been discussed. The suggested model uses statistics analysis, such as Z-Score, alongside machine learning algorithms, including Isolation Forest, to detect any abnormalities in the credit card transactions.

As stated above, the suggested methodology helps detect abnormalities in the transactions without the requirement of labeled fraud information, thereby rendering it appropriate for actual scenarios where fraud cases are scarce.

This proposed model streamlines the process of fraud detection and helps financial analysts identify potential fraudulent transactions more efficiently. Further research on this topic could explore the use of more machine learning algorithms and real-time transaction monitoring alongside deep learning methods.

V. ACKNOWLEDGMENT

The authors would like to extend their heartfelt thanks to the teaching staff of the Department of Computer Science and Engineering for their help and cooperation during the entire process of developing this research. Their constructive advice played an instrumental role in bringing this research to fruition.

REFERENCES

[1]. Credit Card Fraud Detection Dataset, Kaggle.
[2]. Scikit-learn: Machine Learning in Python, by Pedregosa et al., Journal of Machine Learning Research.
[3]. Isolation Forest, by Liu, Ting; Zhou, IEEE International Conference on Data Mining.
[4]. Anomaly Detection: A Survey, by Chandola, Banerjee, Kumar, ACM Computing Surveys.