CREDIT CARD FRAUD DETECTION SYSTEM USING ADVANCED BIDIRECTIONAL GATED RECURRENT UNIT

Revoori Swetha
Assistant Professor, Department of Computer Science and Engineering, KITSW, Warangal, Telangana, India.

Enugala Rajitha
Assistant Professor, Department of Computer Science and Engineering, KITSW, Warangal, Telangana, India.

Manda Niharika
Assistant Professor, Department of Computer Science and Engineering, KITSW, Warangal, Telangana, India.

Masani Ruchi Nandan
Department of Computer Science and Engineering, Vaagdevi College of Engineering, Bollikunta, Telangana, India.

ABSTRACT

Credit card fraud refers to the physical loss of credit card or loss of sensible credit card data. Several machine-learning algorithms can be applied for detection of false credit card activities. Financial fraud is an ever-growing threat, with real results in the business activity. Machine learning had performed an essential role in the discovery of credit card fraud in online transactions. The performance of fraud detection in credit card transactions is influenced by the sampling method on the dataset, collection of variables and detection technique(s) applied. Consequently, applications of detecting credit card frauds are increasing for high-value banks and financial institutions on demand. False activities can happen in many ways and can place into several categories. Financial fraud, such as money laundering, is a severe process of crime that makes illegitimately obtained funds go to terrorism or other criminal activity. The primary issue when it happens to represent fraud detection as a classification difficulty comes from the reality that in real-world data, the majority of transactions are not false. This variety of unauthorised action requires complex networks of business and financial transactions, which perform it challenging to detect fraud entities and find the characteristics of fraud. In this paper, the class imbalance problem is handled by finding legal or fraud transaction using advanced bidirectional Gated recurrent unit.
Credit Card Fraud Detection System Using Advanced Bidirectional Gated Recurrent Unit

(ABiGRU) based machine learning algorithm. Also, suggesting advanced frequent pattern mining algorithm. It can leverage both network data and function data for the detection of financial fraud and very opportunity presented using the best machine learning paradigm. The experimental results illustrate that the proposed scheme provides better accuracy compared with the previous algorithms.

Keywords: Credit card fraud detection, anomaly detection, applications of machine learning, Machine learning


http://www.iaeme.com/IJARET/issues.asp?JType=IJARET&VType=11&IType=12

1. INTRODUCTION

Credit cards are commonly managed due to the primary transfer of e-commerce and the development of smart cellular devices. Non-existent card transactions (CNP) (i.e. online transactions without physical card) [1] are additional and unusual, especially because all credit card transactions are performed via portals on the network, along with PayPal and Alipay. Online credit card transactions are becoming more comfortable and more beneficial. However, there may be an increasing heresy in transaction fraud that ends in huge financial losses every year. The losses were expected to grow annually through double-digit fees by 2022 [2]. Since the body card is not necessary for the online transaction environment, and the cardboard records are sufficient to complete the payment, the fraud is much less complicated than before. Transaction fraud has become a widespread obstacle to improving e-commerce and has a widespread impact on the financial system. Therefore, fraud detection is necessary and important.

Fraud detection is the method of tracking cardholder transaction behaviour to see if there is an incoming transaction by the cardholder or others [3]. In general, there are two types of fraud detection techniques: abuse detection and anomaly detection. Detection of category abuse is used to determine whether or not the transaction is a scam. In general, this method should identify the types of fraud that exist for the model by knowing the distinctive fraud patterns. Anomalous disclosure of the creation of a cardholder profile includes the daily behaviour of the transaction, entirely based on beyond the transaction statistics, and the identification of the new transaction as a capacity fraud if it deviates from the normal behaviour of the transaction. However, the discrepancy detection approach requires sufficient sequential data to distinguish between the cardholder’s daily transaction behaviour.

Fraud increased dramatically with the advancement of advanced technology and global communications. [4] Fraud can be prevented in two basic ways: prevention and detection. Prevention Avoid any fraud attack by acting as a security layer. Detected after the block has already failed. Therefore, detection simplifies identification and alerting once a fraudulent transaction is initiated. Recently, non-existent card transactions [5] in credit card operations have validated the popularity of online price portals. According to a Nelson report released in October 2016, online pricing systems generated more than $ 31 billion worldwide in 2015, up 7% from 2014. Global losses from credit card fraud accelerated to $ 21 billion in 2015 and will reach $ 31 billion by 2022; but there was a boom in fraudulent transactions that greatly affected the economic system. Credit card fraud can be classified into several categories. There are basic types of fraud that can be effectively identified in a variety of transactions: Card Non-Exist (CNP) and Card Presence (CP). These types can be described in more detail by disaster money fraud, theft / fraud, fraud, and behavioral fraud in many packages.
2. LITERATURE SURVEY

In early investigations, several methods have been suggested to transfer responses to identify fraud from supervised procedures, and non-controlled methods to hybrids. It makes it necessary to examine techniques related to the detection of credit card fraud and have explicit experience in forms of credit card fraud. Over time, fraud techniques that added new fraud strategies and making them an enthusiastic area for detective fraud activities. The remainder of this section describes a device for analyzing algorithms, model study tools and fraud detection systems that were used to detect fraud. Problems arose in the survey for the subsequent use of operating an environmental device to obtain knowledge of the copy.

Maniraj et al. [6] Credit card companies need to be aware of fraudulent credit card transactions so that customers are not charged for devices they no longer purchase. These issues can be addressed using data science, and their importance alongside studying the system cannot be overstated. The reason for this company is to show modeling a range of statistics on device usage, and to gain knowledge when credit card fraud is detected. The credit card fraud detection problem is a model that goes beyond credit card transactions with the facts for those who have become fraud. This release is used to understand whether the new transaction is fraudulent or not. Our goal here is to obtain 100% of fraudulent transactions and reduce false fraud results. Detecting credit card fraud is popular because of its beauty. In this generation, we focus on statistical hardware for reading and pre-processing, as well as implementing anomaly detection algorithms in conjunction with the Outlier Factor and Isolatation Forest neighborhood base of credit card transaction records converted by the PCA.

Abhimanyu et al. [7] Credit card fraud caused a $3 billion deficit in North American monetary institutions in 2017. Increasing pressure from virtual collections systems such as Apple Pay, Android Pay and Venmo resulted in a loss due to the fraudulent interest transfer. Deep learning has added a promising era to the hassle of detecting credit card fraud by using licensed organizations to take full advantage of outdated consumer data as well as real-time transaction statistics recorded at the time of the transaction. In 2017, a set of insights shifting the insight gained from the era provided similar results to the prevailing fraud detection strategies with slow and logistical degradation. However, the in-depth analysis covers many systems.

Moreover, the different parameters used for transcription clustering (for example, the population of neurons in the hidden layer of the nervous community) also influence their outcomes. The authors evaluated a subsection of a deep analysis, from the known artificial neural community to topology with embedded time and memory additions, which consist of short memories and long periods of time, and various criteria while demonstrating their effectiveness in fraud detection. In a set of records of nearly eighty million transactions, credit card use was previously identified as fraudulent and legitimate. She saw the performance-saving overall computing environment as accelerating the cloud to bypass unusual fraud detection issues, as well as massive disruption and scalability. This assessment provided a comprehensive guide for assessing the sensitivity of implementation standards with an indication of standard fraud detection performance. In addition, it has provided a framework for improving the in-depth study of credit card fraud attacks and enabling monetary institutions to reduce losses with the help of the use of fraudulent sports.

Akila et al.[8] Credit card fraud customers take billions of invoices in a year. Even with multiple systems in the region correct fraud detection remains unresolved due to the many underlying issues listed in transaction logs. Here authors have discussed the essential nature of the studies. They have proposed a package to launch in particular on the risks of the Ukrainian Reserve Bank as a technique to address the main problems presented in the events, in addition to providing realistic effects. The traditional padding version has been raised, and new word
updates have been protected from the ongoing and value-sensitive series. Bag filling operations are powerful in managing unbalanced data. At the same time, the inclusion of Naïve Bayes, which relies particularly on risk, addresses the underlying noise that was provided in the transaction information. Cost-sensitive fusion replaces traditional fusion of voting to provide results that show extreme performance and occasional pricing. Comparisons with current models indicate excessive-performance ranges for the proposed version of RBE.

Dilip Singh et al. [9] the branch of online transactions has significantly collapsed with each passing day. Credit card transactions are a problem with these transactions. Economic losses rose similarly, along with credit card fraud. Therefore, fraud detection systems have been of great importance to banks and monetary institutions. Since the frauds were not now evaluated in standard transactions, they arose with the discomfort of an imbalance of beauty, and to deal with this imbalance, modelling techniques were used. Unsupervised samples (SMOTE, SMOTE ENN, SAFE SMOTE, ROS, and SMOTE TL) were also performed. In Instance Information, we apply touch pix (CSVM, C4.Five) and Taxonomy Project (Adaboost, Bagging) to evaluate the overall performance of using Sensitivity, Privacy, G Concept, and Location under ROC. We understand that the SMOTE ENN era detects fraud higher than specific classifiers in the set of incremental sampling strategies and that TL performs higher within the final set of sampling techniques that I have taken.

Jarrod West et al. [10] financial fraud has proven to be a widespread problem at some point in history. Because of its significant impact on society’s view of the gentle way to solve it, it is essential. Various experimental problems may be related to the detection of complete financial fraud based on arithmetic intelligence. In this paper, they could discuss 3 of them: the general rule of willingness to disclose, standard performance metrics, and choice of features. The characteristics of these three problems are described as being understood in contemporary literature. However, we found that there are widespread research gaps that require additional attention. We will reduce this deficit by researching extensively in detection algorithms, standard performance metrics and feature selection, using a series of controlled simulations of the problem of credit card fraud and assessing the consequences.

Fahimeh et al. [11] Due to the rapid rise in e-commerce and virtual imaging systems, credit card fraud has accelerated. The reason for this file is increased credit card fraud detection (CCFD) that specifically relies on Artificial Neural Networks (ANN) and is uselessly priced to reduce threat recognition and the chance of loss. ANN's strategy was used to bail him out and detect credit card fraud. Due to the unbalanced nature of statistics (fraud and non-fraud cases), fraudulent transactions are difficult to achieve. To do this, the problem of unbalanced sports, the over-priced machine, became a source. The proposed model, known as a cost-sensitive neural network (CSNN), relies specifically on generating abuse discovery. Compared to the full version based on AIS, this one showed financial savings in discovery fees and double fees. Records for this analysis are taken from actual transaction statistics provided by a large Brazilian credit card company.

Zareapoor et al. [12] Credit card fraud Credit card fraud was increased with the growth of the current generation and the global ways of conversation. Credit card fraud bills paid billions of dollars to the economic organization annually, and fraudsters were continually tried to discover new rules and methods for committed illegal actions. Therefore, fraud detection systems have proven vital to banks and the monetary organization to reduce their losses. However, there could be a lack of published posts about credit card fraud detection because investigators should not maintain a credit score card transaction data set. The best widely used strategies are Naïve Bayes (NB) fraud detection methods, Support Vector Machine (SVM), and K-Nearest Neighbourhood (KNN). These strategies could be used with a combination of standard techniques or meta-study to build compilations. But among all the evolving strategies,
business identification techniques are defined as famous and not unusual, not now for particularly direct execution, but also for their excellent overall predictive performance on sensitive issues. we train several document extraction techniques used to detect credit card fraud and examine every technology that mainly depends on the exact design criteria. After many checks and comparisons, we introduced the third choice filler class as a successful classifier to build a fraud detection copy. The overall performance assessment within the real-time credit card transaction dataset was completed to demonstrate gains in the group's guideline group.

Halvaeie et al. [13] a wide range of transactions was evolving online today. A significant problem with these transactions involves credit card transactions. On the other hand, the high rate of online fraud is fantastic, and usually, the result of generation reaches for all people. Studies have been conducted on many tools and techniques for saving money and detecting credit card fraud. The artificial immune system is, without the uncertainty of them. However, companies need precision along with internal rhythm fraud detection structures, which are fully acquired. In this document, we cover credit card fraud detection, the use of artificial immune systems (AIS), and introduce a new model known as the Fully AIS Based Fraud Detection Mechanism (AFDM). We will use an AIRS-stimulated algorithm and improve it to attack fraud. Increased accuracy to 25%, reduced rate to 80%, and reduced system reaction time to 40% compared to the basic algorithm.

3. PROPOSED METHODOLOGY

In this paper, we like to extend a new framework of credit card fraud detection by considering the anomaly detection method of the private location of entities, and fraudulent behavior. Specifically, we check: (1) how to make use of distinct matrix to detect and track fraud; (2) how to make a mathematical copy of each graphical matrix and matrix of functions that will simultaneously fulfill fraud detection and tracking obligations. To clarify these problematic situations, we have proposed a unique framework for Advanced Fraud detection framework, as shown in Figure 1, for economic activities, especially money laundering activities. We incorporate fraud detection and anomaly function detection into the corresponding framework for selecting comparable fraud patterns and functions at the same time. The combination of entity discovery and feature discovery enables us to create a new fraud detection framework to generate robust and sparse financial reporting. In essence, applied fraud patterns assist in identifying fraud, and related features help expose the nature of fraud activities.

Our pilot study of artificial and real global registry units demonstrates the efficacy of Proposed method, which detects a fraud pattern and determines fraud-related characteristics in an uncensored manner by searching for low-range, approximate representations, and matrix residues. The main participants of the paper can be paraphrased as follows.

- Implement a proposal work to create Legal/Fraud Pattern
- Demonstrate unique cases of cash fraud and formulate fraud patterns in terms of the chart.
- proposing a unique, advanced frequent pattern mining algorithm, for the problem of discovering intricate patterns and identifying anomaly objectives, and using the residual evaluation of two arrays in a society based entirely on the economic graph;
- Evaluate the framework using artificial and real data from around the world to demonstrate the effectiveness and performance of the proposed set of rules.
3.1. Fraud detection using proposed BiLSTM Mechanism

The model carries an embedding of the specific characteristics and 0.1 of the output of the embedding has been eliminated to avoid excessive modification with the help of using a spatial leakage layer. The result is then converted to a dimension through the flat layer and then fed into the spill and leakage layer. Appropriately, the output was delivered to the BiLSTM and BiGRU layers simultaneously, as the maximum global aggregation was applied to the outputs of each model so that it extracted the most important maximum characteristics and then the output was combined.

When the proposed algorithm is applied to a customer's credit card transaction data, it returns a set of attributes that display the same values in the set of transactions defined by the support.

The Matching algorithm (test) is explained below,

**Step 1.** Count the number of attributes in the incoming transaction matching with that of the legal pattern of the corresponding customer. Let it be lc

**Step 2.** Count the number of attributes in the incoming transaction matching with that of the fraud pattern of the corresponding customer. Let it be fc.

**Step 3.** If fc = 0 and lc is more than the user defined matching percentage, then the incoming transaction is legal.

**Step 4.** If lc = 0 and fc is more than the user defined matching percentage, then the incoming transaction is fraud

**Step 5.** If both fc and lc are greater than zero and fc ≥ lc, then the incoming transaction is fraud or else it is legal.

The pseudo code of the testing algorithm is given below,

3.2. Proposed Advanced Frequent Pattern Mining Algorithm

**Input:** LPD is Legal Pattern Database, FPD Fraud Pattern Database, Incoming Transaction T, Number of Customers "n" Number attributes "k" matching percentage "mp"

**Output:** 0 (if legal) or 1 (if fraud)

**Assumption:**

First attribute of each record in pattern databases and incoming transaction is Customer ID.

If an attribute missing in the frequent item set then we considered it as invalid.

**Begin**

lc = 0; //legal attribute match count
fc = 0; //fraud attribute match count

for i = 1 to n do

if(LP(i, 1) = T(1)) then //first attribute
for j = 2 to k do
    if (LPD(i, j) is valid and LPD(i, j) = T(j)) then
        lc = lc + 1
    endif
endfor
endif
endfor
for i = 1 to n do
    if (FPD(i, 1) = T(1)) then // first attribute
        for j = 2 to k do
            if (FPD(i, j) is valid and FPD(i, j) = T(j)) then
                fc = fc + 1
            endif
        endfor
    endif
endfor
endif
endif
for i = 1 to n do
    if (fc = 0) then // no fraud pattern
        if (lc/no. of valid attributes in legal pattern) ≥ mp) then
            return (0); // legal transaction
        else return (1); // fraud transaction
    endif
    elseif (lc = 0) then // no legal pattern
        if (fc/no. of valid attributes in fraud pattern) ≥ mp) then
            return (1); // fraud transaction
        else return (0); // legal transaction
    endif
After finding fraud patterns and legal patterns for every user, the fraud detection mechanism passes through fraud databases and criminal fraud detection patterns. These heresy databases are much smaller than the original customer transaction databases because they only contain the consumer's record. This research proposes an identical algorithm that covers sample databases to maintain the validity of the incoming transaction to detect fraud. If the request closest to the legal pattern of the corresponding pattern is found, the matching algorithm returns "0", and this gives a new signal to the bank to allow the transaction. If the closest match to the consumer fraud sample is observed, the matching algorithm returns "1", to alert the financial institution to avoid the transaction. The length of the pattern databases is $n \times k$, where $n$ is the number of clients and $k$ is a variety of attributes.

### 3.3. Real time Fraud Detection

In the afterlife, fraud was revealed with the help of group transactions and the use of instrument learning techniques in them. Since the effects can be seen after weeks or months, monitoring for detected fraud becomes very difficult, and there have been many cases where fraudsters control the dedication of many fraudulent purchases before they are identified. Real-time fraud detection is the implementation of innovative fraud detection within seconds online. This way, our devices can detect real-time fraud. It provides an alert to the economic establishment showing the fraud model and rate of accuracy, making it difficult for fraud tracking organizations to move forward with their subsequent actions without having to waste their time and money.

### 3.4. Fraud Detection System

Immediate exposure of credit card fraud can be considered a significant contribution to this allocation. The real-time fraud detector consists of three devices number one; API Unit, Fraud Detection Bureaucracy and Record Store. All plugins are concerned with fraud detection at the same time. Fraudulent transactions are categorized into four types of fraud (MCC-generated serious fraud, ISO response code, unrecognized net transaction, and a transaction over $100) using three supervised understandings. The API unit is responsible for switching operations in real-time between the fraud detection form, the graphical user interface and the store of facts. The record store is used to save the expected results of residence factors and various related tooltips to take advantage of the fashion vision. A person can interact with the fraud detection tool through graphical consumer interfaces that advise real-time transactions, fraud indicators, and old fraud events on the chart. When the transaction is identified as fraudulent using a form...
of fraud detection, a message can be sent to the API unit. The API unit then informs the detained person with the assistance of sending a notification and comments.

3.5. System Architecture

![Proposed System Model](image)

**Figure 1. Proposed System Model**

4. RESULTS AND DISCUSSIONS

The output of the metrics depends on the results obtained by True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN).

- **True Positive (TP):** The transaction cases which are not fraud and the system model has predicted as not fraud.
- **True Negative (TN):** The transaction cases which are fraud and the system model has predicted as fraud.
- **False Positive (FP):** The transaction cases which are fraud and the system model has predicted as not fraud.
- **True Negative (TN):** The transaction cases which are not fraud and the system model has predicted as fraud.

**Evaluation Metrics**

We employed various metrics because the dataset utilised in this paper was highly imbalanced, using the accuracy metric individually will not be accurate to measure the Accuracy of the approach. Here we are using three metrics of accuracy, sensitivity, and specificity.

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FP}
\]

\[
\text{Specificity} = \frac{TN}{TN + FP}
\]

To evaluate performance, we primarily relied on the location below the ROC curve to evaluate results.
Table 1. ROC for the obtained readings

<table>
<thead>
<tr>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>FPR</th>
<th>TPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>25</td>
<td>30</td>
<td>12</td>
<td>0.454545</td>
<td>0.733333</td>
</tr>
<tr>
<td>127</td>
<td>187</td>
<td>102</td>
<td>84</td>
<td>0.647059</td>
<td>0.601896</td>
</tr>
<tr>
<td>411</td>
<td>187</td>
<td>275</td>
<td>160</td>
<td>0.404762</td>
<td>0.71979</td>
</tr>
<tr>
<td>623</td>
<td>285</td>
<td>408</td>
<td>184</td>
<td>0.411255</td>
<td>0.771995</td>
</tr>
<tr>
<td>749</td>
<td>392</td>
<td>551</td>
<td>308</td>
<td>0.415695</td>
<td>0.708609</td>
</tr>
<tr>
<td>867</td>
<td>448</td>
<td>819</td>
<td>366</td>
<td>0.353591</td>
<td>0.703163</td>
</tr>
<tr>
<td>1034</td>
<td>522</td>
<td>1093</td>
<td>351</td>
<td>0.32322</td>
<td>0.74657</td>
</tr>
</tbody>
</table>

Also table shows the TPR and FPR readings.

Figure 2. ROC proposed fraud detection framework

As shown in Figure 2, samples from the fraud detection framework proposed by the ROC are used with exceptional magnificence in this work. Increasing the amount of negative exercise will not affect the result. All various samples expect the same accuracy. All the suggested measures can be determined from the confusion matrix.
Figure 3 Comparison of time with different Rank size r

We evaluated the temporal illustration of the proposed approach, RPCA and SVD with a variety of ranges, r to measure the closure matrix. The result is illustrated in Figure 3: The suggested framework benefits excessive time performance.

Figure 4. Feature matrix from synthetic data

The Detection of an anomaly in the use of subspace aggregation based on the belief that they grouped in the subspace anomalies in the method of small samples. The next test is to evaluate the proposed method using three techniques MAFIA, SCHISM, and DISH, which have an excellent performance of aggregation in subspace. We comply with the proposed standards for three aggregation strategies. We exchange a variety of anomalies to detect and archive quality. From Figure 4, we see that the proposed technique achieves high detection accuracy in a matrix similar to synthetic records and real international information.

5. CONCLUSION

Detecting credit card fraud has been a particular field of investigation for investigators for years and will be an exciting area to investigate soon. It occurs primarily because of the constant change in fraud patterns. In this paper, we recommend a new advanced fraud detection system for credit card transactions by discovering four different types of fraudulent transactions using the best adaptation algorithms and addressing related issues identified by previous fraud
Credit Card Fraud Detection System Using Advanced Bidirectional Gated Recurrent Unit

detection investigators—credit cards. By processing credit card fraud detection in real-time using predictive analytics and the API unit, the end-user notified through the graphical user interface per second in which fraud occurs. With this suggested framework, administrators in the financial direction can not only discover the fraud patterns but also determine the original of fraud with different feature. Financial activities are time-dependent. We can represent the competition in tensor similarity and feature tensor. So we'd like to take a look at how to integrate the tensor in Fraud Detection Framework. This section of the system can provide the fraud research organisation to decide to proceed to the next step once a different transaction discovered. The proposed method shows a comparison of time with different Rank size ‘r’ for different methods. The experimental results showed that the proposed method has better accuracy compared to previous methods.

REFERENCES


