Review

Machine Learning-Based Detection of Credit Card Fraud: A Comparative Study

Zainab Khamees Alkhateeb and Abeer Tariq Maolood

Department of Computer Science, University of Technology/Baghdad, Baghdad, Iraq

Abstract: One of the fastest-growing problems with a high impact on the financial sector is financial fraud. Recently, data mining has been identified as one of the effective ways of detecting fraudulent credit card transactions. As a data mining problem, the detection of fraudulent credit card transaction is a challenging task due to the following reasons: (i) The frequent changes in the patterns of normal and fraudulent activities and (ii) the high level of skewness related with credit card fraud datasets. The aim of this article is to review the existing techniques for fraudulent transactions detection in credit cards, with more focus on the techniques that are Machine Learning (ML) based and nature inspired-based. The recent trend in the detection of credit card fraud was also presented in this article. Furthermore, the limitations and usefulness of the existing techniques for fraudulent transaction detection in credit cards were also outlined. The necessary fundamental information for further studies in this area was also provided. This review will also guide individuals and financial institutions seeking for effective techniques for credit card fraud detection, especially those that are based on ML and nature-inspired algorithms.

Keywords: Credit Card, Fraud Detection

Introduction

The progression of the existing technology and worldwide communication has resulted in an increased rate of fraudulent activities (Halvaei and Akbari, 2014). Fraud can be curbed by either preventing or detecting its occurrence.

Data prevention involves the formation of a protective layer around the data to prevent external attacks. The aim is to prevent the occurrence of fraudulent activities on the data. Contrarily, fraud detection involves the identification of fraudulent activity and triggering the required response as soon as the activity perpetrated. This implies that detection is the second line of defense (triggered when prevention has failed). It is, therefore, important to ensure that detection is always enabled since it may not be possible to predict when a given protection technique will fail (Michael and Pedro, 2009; Adrian, 2015). Financial fraud is a critical problem in corporate and finance business as it affects several economies, businesses and cost of living. The different types of fraud are shown in Fig. 1. The pattern and characteristics of normal and suspicious financial transactions can be determined using data processing techniques supported by expert knowledge of normal and abnormal behaviors (Shukur, 2019).

Credit Card Fraud

This is one of the major types of frauds with significant importance in the banking sector. Hence, there is a need to strengthen the existing techniques for fraud detection with security systems that aim at fraud prevention. A system will perform well if the fraud detection system is fast in its responsibility. Card transaction is applicable to both online and regular purchases; hence, the pain of a fraudulent credit card transaction is felt by both the shoppers and the merchants as they are both subjected to economic loss (Mahmoudi and Duman, 2015). This is an important issue that requires both the issuing banks and the card manufacturers to solve by investing mainly on its prevention (Halvaei and Akbari, 2014). Although online shopping and payment platforms ensure convenient, comfortable, easy and seamless payment of goods and services, there are still issues of financial losses associated with e-commerce which cannot be ignored. Coping with these problems require the banks and organizations to deploy good security techniques which can adapt to the changes in the nature of fraudulent activities with time. Credit card transaction can be done either physically or virtually.
The physical method required that the card must be used to make a swipe, but for the virtual method, the transaction is approved by providing some card details, such as the CVV number, the name of the cardholder, the security question, the password, etc., (Zareapoor et al., 2012). Fraud can either be prevented or detected as it occurs. Fraud prevention aims at preventing the occurrence of fraudulent activity; such transactions are spotted and denied authorization (Sachin and Duman, 2011). For fraud detection, the major target is to distinguish normal activities from fraudulent ones (Quah and Shrioganesh, 2008).

**Securities or Commodities Fraud**

This refers to the several techniques used by a fraudster to deceive a person based on false information to invest in a company. Such methods include the Ponzi and Pyramid Schemes, Hedge Fund Fraud, Embezzlement and Foreign Exchange Fraud (West and Bhattacharya, 2016).

**Financial Statement Fraud**

A financial statement is an official company document that details their financial status in terms of their income, expenses, loans and profits. Such documents can be used to show the status of a company to influence stock prices. Financial statement fraud (or corporate fraud) refers to the fraudulent manipulation of the financial status of a company in a bid to evade taxation, improve stock performance, or exaggerate performance as a result of managerial pressure (West and Bhattacharya, 2016). It may be difficult to detect financial statement fraud when the basic understanding of the sector is lacking. Another factor that makes it difficult to detect is that it is perpetrated by experts in the field who can easily cover their fraudulent activity (Sahin and Duman, 2011a).

**Insurance Fraud**

This refers to any type of fraud that is associated with any step of an insurance process and committed by the people in the sector. It is encountered when the fraudulent user submits an insurance claim that is exaggerated. This type of fraud comes in the form of excessive billing, kickbacks and duplicate claims (West and Bhattacharya, 2016).

**Mortgage Fraud**

This is a special form of financial fraud in which a mortgage document or property is manipulated with the aim of misrepresenting the actual value of the property or document just to influence the funding of the property loan by the lender (West and Bhattacharya, 2016).

**Money Laundering**

This is an act committed by criminals when trying to invest the proceeds of illicit activities into valid ventures with the aim of hiding the original source of the money and appearing legitimate just to deceive the appropriate authorities from tracking their crimes. Money laundering is so dangerous that it will raise the economic influence of the criminal (West and Bhattacharya, 2016).

**Related Works**

Several works have been reported on the detection and prevention of credit card frauds. For instance, an approach for credit card fraud detection which combined SVM with decision tree was proposed by Sahin and Duman (2011a). The study evaluated the performance of...
different DT-based classifiers and several variants of SVM. A fuzzy clustering and neural network-based approach were proposed by (Behera and Panigrahi, 2015) for fraudulent baking transactions detection. This approach detects banking fraud in three phases; first, the user and his card details are authenticated and verified, followed by a performance of fuzzy means clustering to determine the normal usage pattern of the user based on his previous transaction history. Upon the detection of a new but doubtful transaction, the NN mechanism will be applied to the nature of the new transaction (whether fraudulent or genuine). Another study by (Ng and Jordan, 2002) compared Naïve Bayes (NB) and Logistic Regression (LR) for fraud detection (Ng and Jordan, 2002). From the analysis, it was shown that despite the lower asymptomatic error of the discriminative LB algorithm, the generative NB classifier may rapidly converge to its higher asymptotic error. Some studies have reported better performance of LR compared to NB; however, this is based on small datasets.

**Existing Techniques**

Several computational and statistical data mining techniques exist. This section outlined the role of the existing methodologies in the literature:

**a. Support Vector Machine (SVM)**

SVMs are classifiers which label and classify data in the feature space. They are applicable to linear (separable and inseparable) datasets. In a linearly separable dataset, a straight line can be used to demarcate the data of class A from that of class B. For the linearly inseparable data, it is not possible to identify the linear line that will maximize the data classification. The SVM mainly aims at mapping a hyperplane which will cluster the data vectors into clusters. The linearly separable dataset may have several hyperplanes, but the task is to identify the best hyperplane that will guarantee the maximum inter-class margin. For instance, a binary dataset that has \( x \) as the hyperplane will amount to the following definitions:

\[
G_{i,j} = \frac{1}{V_i \in \text{class}1} \geq 1, \quad G_{i,j} = \frac{-1}{V_i \in \text{class}2}
\]

The support vectors are the points lying on the boundaries; they define the hyperplanes. Most classifiers execute linear classifications by creating a linear line within the feature space. Nonlinear data classification can also be performed by extending the linear data classifier using the following steps:

**Step 1.** The original data should be transformed in a manner that will map it into a high dimensional space

**Step 2.** The hyperplane that will provide the best classification of the data in the new high dimensional space should be searched

**b. Neural Network (NN)**

A neural network mimics the biological function of the human brain. It was developed as a computational representation of the human nervous system where synapse and neurons are represented using a graph of edges and vertices (Ng et al., 2011). Figure 2 showed the input variables modeling in an NN as a layer of vertices. Every connection in the graph is assigned a weight function. The other vertices are assigned to separate levels which portrays “the distance from the input nodes” (Kirkos et al., 2007). Thus, “each nodal input is a function of the vertices connected to the preceding layer”. The signal received per neuron, \( j \), is represented as:

\[
U_j = eW_i \times X_i
\]

Where:

- \( W_i \) = Connection weight between the two neurons (\( i \) and \( j \)) and
- \( X_i \) = The input of neuron \( i \)

Should the outcome of this representation be higher than an already predefined limit, the present neuron will “fires” and become the next layers’ input (Kirkos et al., 2007).

**c. Artificial Immune System (AIS)**

The AIS is a form of DM technique which depends on the concept of the natural immune system to discover antigens (Sx and Banzhaf, 2008). The AIS can be used to simulate several biological behaviors; however, some of the AIS-based models can only create the detector cells based on their foreign body’s detection capability. The detector cells are randomly generated, after which simulation is executed to evaluate the algorithmic effectiveness in terms of the training performed by the various classification techniques. The 2 common variants of AIS are “Clone Selection (CS) and Negative Selection (NS)”. Regarding CS, the generated detector cells live a short life but if they are able to detect an antigen within their short life, their life will be prolonged so that they can fight off the antigen. At the end of the fight with the antigen, the CS will mutate and the ones that are best suited for the detection of the antigens at the end of the simulation are called the survival cells. For the NS, detector cell creation is done arbitrarily and from the created cells, the one that will react with intruders will be selected from the whole system while the rest will be discarded (Halvaeie and Akbari, 2014).
d. Bayesian Belief Network (BBN)

The BBN is a statistical method of classifying problems which depends on the Bayes theorem and work on the concept of establishing the chances of an hypothesis being true (Ngai et al., 2011). Given the hypothesis H for the study, the probability P is determined as follows:

\[ Y = \begin{cases} 0 \\ 1 \end{cases} \]

The calculation of the probability of sample xi being a member of class one is done as follows:

\[ P\left( Y_i = 1 \mid x_i \right) = \frac{\exp\left( w_0 + w^T x_i \right)}{1 + \exp\left( w_0 + w^T x_i \right)} \]

Where:
- \( W_0 \) = The intercept
- \( W \) = The coefficient vector

Both are regression standardization parameters (Ravisankar et al., 2011).

e. Logistic Regression (LR)

The LR is a statistical-based binary classification method which utilizes a linear model (Ngai et al., 2011) to perform regression on a set of variables (Ravisankar et al., 2011). The LR is commonly used for the prediction of the patterns of a dataset with numeric or unambiguous attributes (refer to Fig. 3) (Ngai et al., 2011). It uses the logarithm to computes probability from several input variables and one response dependent variable:

\[ Y = \begin{cases} 0 \\ 1 \end{cases} \]

f. Decision Tree (DT)

The DT uses a combination of binary trees and nodes during data classification (refer to Fig. 4). When s sample moves along the tree, the nodes belonging to such sample will be generated. Then, the tree is partitioned into subsets and later stored in the “mutually exclusive subgroups” (Kirkos et al., 2007); hence, it is referred to as “classification and regression tree” (Kirkos et al., 2007).

Another method known as pruning has also been suggested to address the issue of overfitting (Sahin and Duman, 2011b). With pruning, the tree nodes can be removed without affecting the general models’ accuracy.

Fig. 2: Modeling of input variables in NN
g. **Self-Organizing Map (SOM)**

The SOM is like ANN as both are comprised of one matrix of neurons. This technique uses a non-linear algorithmic framework to transform the input variable into a 2-D array, with the primary aim of modeling similar input variables which are nearer to the target matrix as neurons and provide a view of the input. Then, various distance function (such as Gaussian formula, Euclidean distance formula) is applied on the set of nodes (Halvaeie and Akbari, 2014). A clustering function is applied to each neuron; the clustering function is given by:

\[ Y_{i+1} = Y_i + \alpha (X_i - Y_i - 1) \]

Where:
- \( Y_i \) = A specific nodes’ current weight
- \( X_i \) = The present input vector
- \( \alpha \) = A distance-related function

Before terminating the algorithm, clustering must be performed on a set of iterations (Olszewski, 2014).
h. Hybrid Methods

The hybrid methods are developed for specific types of problems. It is a hybridization of more than two similar (in terms of benefits) conventional methods to generate a stronger algorithm. There are several ways of building hybrid models, such as the highest-level technique and lower level (preprocessing stage) technique. In the highest-level technique, linearity is applied; the output of the first stage is the input for the next stage (Duman and Ozcelik, 2011). The individual steps of the conventional algorithm may combine in the hybrid model to build an entirely new step or system (Duman and Ozcelik, 2011). Hybrid models are used for specific problem domains where the target is to achieve a different aspect of performance such as computation efficiency, classification ability and ease of use. Regarding the lower level (pre-processing step) technique, data modification is first performed prior to classification (Jans et al., 2011). Table 1 presents a summary of the strengths and limitation of the reviewed existing methodologies.

Table 1: Summary of the strengths and limitation of the existing methodologies

<table>
<thead>
<tr>
<th>Method</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural network</td>
<td>- It is a recognized fraud detection method which can be applied to several binary classification (non-algorithmic) problems.</td>
<td>- The training and operation demand high computational power; hence, it is not ideal for real-time application.</td>
</tr>
<tr>
<td>Logistic model</td>
<td>- Easy to use and well known for fraud detection.</td>
<td>- It may overfit if the training dataset does not represent the problem space perfectly, thereby requiring constant updating to be suitable for new types of fraud.</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>- It can handle non-linear classification issues such as fraud detection.</td>
<td>- It has a low classification performance compared to the other data mining techniques and has trouble with fraud detection complexity.</td>
</tr>
<tr>
<td>Decision trees</td>
<td>- Easy to use and understand.</td>
<td>- It may overfit if the training dataset does not represent the problem space perfectly, thereby requiring constant updating to be suitable for new types of fraud.</td>
</tr>
<tr>
<td>- It requires low computational power and minimal training, thereby suitable for real-time application.</td>
<td>- REQUIRES high computational power to optimize the initial setup.</td>
<td></td>
</tr>
<tr>
<td>Bayesian belief network</td>
<td>- Suitable for other binary classification (non-algorithmic) problems.</td>
<td>- The knowledge of normal and abnormal patterns is needed to investigate fraud.</td>
</tr>
<tr>
<td>Genetic algorithm</td>
<td>- Easily implementable using classification accuracy as the fitness functions.</td>
<td>- The training and operation demand high computational power; hence, it is not ideal for real-time application.</td>
</tr>
<tr>
<td>- Suitable for other binary classification (non-algorithmic) problems.</td>
<td>- The issue of local maxima/minima makes it difficult to adapt to new types of fraud.</td>
<td></td>
</tr>
<tr>
<td>Self-organizing map</td>
<td>- Easy to implement.</td>
<td>- Visualization cannot be easily automated; hence, requires manual observation by the auditor.</td>
</tr>
<tr>
<td>- The visual nature of the results can be easily understood by auditors.</td>
<td>- Its operation demands intensive computational input; hence, it is not ideal for real-time application.</td>
<td></td>
</tr>
<tr>
<td>AIS</td>
<td>- Suitable for tasks that are associated with data imbalance, for instance, fraud detection.</td>
<td>- Being that it may be developed as a new yet-to-be-verified method, it may present a high level of risk considering that fraud detection is a high-cost problem.</td>
</tr>
<tr>
<td>Hybridized methods</td>
<td>- Can easily adapt to new types of fraud as it combined the advantages of several conventional methods.</td>
<td>-</td>
</tr>
</tbody>
</table>
Conclusion

Credit card detection is one of the captivating problem domains. This review points toward ML techniques as the most suitable fraud detection methods due to their high detection rate and accuracy. However, studies are still focusing on improving the accuracy and detection rate of the ML techniques, while organizations are concerned with finding new methods of reducing cost and maximizing profit.

Author’s Contributions

Zainab Khamees Alkhateeb: Writing the manuscript Participated in the experiments, collect the data and literature information, coordinated the data-analysis.

Abeer Tariq Maolood: Contributed to the writing of the manuscript, supervising the manuscript writing.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References


