



Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Detecting credit card fraud by Modified Fisher Discriminant Analysis



Nader Mahmoudi, Ekrem Duman*

Özyeğin University Çekmeköy Campus, Industrial Engineering Department, 34794 Istanbul, Turkey

ARTICLE INFO

Article history:

Available online 6 November 2014

Keywords:

Credit card fraud
 Linear discriminant
 Fisher linear discriminant function
 Modified Fisher Discriminant
 Profitability

ABSTRACT

In parallel to the increase in the number of credit card transactions, the financial losses due to fraud have also increased. Thus, the popularity of credit card fraud detection has been increased both for academicians and banks. Many supervised learning methods were introduced in credit card fraud literature some of which bears quite complex algorithms. As compared to complex algorithms which somehow over-fit the dataset they are built on, one can expect simpler algorithms may show a more robust performance on a range of datasets. Although, linear discriminant functions are less complex classifiers and can work on high-dimensional problems like credit card fraud detection, they did not receive considerable attention so far. This study investigates a linear discriminant, called Fisher Discriminant Function for the first time in credit card fraud detection problem. On the other hand, in this and some other domains, cost of false negatives is very higher than false positives and is different for each transaction. Thus, it is necessary to develop classification methods which are biased toward the most important instances. To cope for this, a Modified Fisher Discriminant Function is proposed in this study which makes the traditional function more sensitive to the important instances. This way, the profit that can be obtained from a fraud/legitimate classifier is maximized. Experimental results confirm that Modified Fisher Discriminant could evenuate more profit.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Nowadays, by increasing credit card transactions in not only online purchases but also regular purchases, credit card fraud is becoming rampant. Today, both merchants and clients are affected in terms of financial losses caused by credit card fraud. Some references reported billions of dollars lost annually due to credit card fraud (Chan, Fan, Prodromidis, & Stolfo, 1999; Chen, Chen, & Lin, 2006). CyberSource (2013) reported in 14th annual online fraud that the actual amount of losses will increase by the increasing online sales. It is also reported that the estimated total loss increased up to \$3.5 billion in 2012 by 30% increase from 2010. Evidently, with the growth in the number of credit card transactions as a payment system, 70% of consumers in U.S. had concerns about identity fraud significantly (McAlearney & Breach, 2008).

Considering this huge amount of financial loss, prevention of credit card frauds is the most concerning issue for researchers in data mining area. Because of large amount of credit card transactions, detecting about 2.5 percent of frauds leads to save over a million dollar per year (Brause, Langsdorf, & Hepp, 1999). However,

along with the development of fraud detection techniques, fraudulent activities done by criminals also have been evolved to avoid detection (Bolton & Hand, 2001). Thus, to perform in the best way, researchers are trying to make modifications in the existing methods or develop new methods to maximize number of frauds detected.

Bolton and Hand (2001) categorized credit card frauds into two groups: application frauds and behavioral frauds. Application frauds occur when fraudsters obtain new cards by presenting false information to issuing companies. On the other hand, behavioral frauds include four types: mail theft, stolen/lost cards, counterfeit cards, and 'card holder not present' fraud. In modern banking system, the more the online transactions increase, the more counterfeit and 'card holder not present' frauds occur; where in both of these two types of fraud, fraudsters obtain credit card details without the knowledge of card holders. Bolton and Hand (2002) presented a good discussion on the issues and challenges in fraud detection research together with Provost (2002).

In the literature, there are many studies made on credit card fraud detection in some of which methods for learning systems are proposed. If we look at these studies, most of the credit card fraud detection systems are using supervised learning algorithms like neural networks (Aihua, Rencheng, & Yaochen, 2007; Juszczak, Adams, Hand, Whitrow, & Weston, 2008; Quah & Sriganesh, 2007;

* Corresponding author.

E-mail addresses: nader.mahmoudi@ozu.edu.tr (N. Mahmoudi), ekrem.duman@ozyegin.edu.tr (E. Duman).

Schindeler, 2006), decision tree techniques such as ID3, C4.5, and C&RT (Chen, Chiu, Huang, & Chen, 2004; Chen, Luo, Liang, & Lee, 2005; Mena, 2003; Wheeler & Aitken, 2000), and support vector machines (SVMs) (Leonard, 1993).

Sahin and Duman (2011) carried out a study using Artificial Neural Network (ANN) and logistic regression (LR) to score transactions where they are flagged as fraudulent or legitimate transactions. They concluded that ANN outperforms LR based on results. However, as skewness of training set increases, the performance of all models decrease.

Aihua et al. (2007) investigated the efficacy of applying classification models to credit card fraud detection problems. Three different classification methods, i.e. decision tree, neural networks and logistic regression are tested for their applicability in fraud detections. Their paper provides a useful framework to choose the best model to recognize the credit card fraud risk based on different performance measures.

In the most of related studies in literature, the cost of a false negative (labeling a fraudulent transaction as legitimate) and a false positive (labeling a legitimate transaction as fraudulent) are taken as equal to each other. However, in this domain the cost of a false negative is much higher than the cost of a false positive and in fact it varies from transaction to transaction. To cope with the higher cost of a false negative, some researches used adjusted cost matrices during the training phase of their classifiers (Langford & Beygelzimer, 2005; Maloof, 2003; Sheng & Ling, 2006; Zhou & Liu, 2006). However, the variable character of misclassification costs is undertaken in only a few studies so far (Duman & Elikucuk, 2013a; Duman & Ozcelik, 2011; Sahin, Bulkan, & Duman, 2013; Sahin & Duman, 2010; Sahin & Duman, 2011).

Actually the main issue in credit card fraud detection modeling is to get the most possible profit from the use of such a classification model. This study, as a pioneer, tries to implement a linear profit based method to maximize total profit where individual benefits and costs of classifying a transaction are considered during the learning phase. That is, the model which is developed is biased towards the correct classification of beneficial transactions than the others.

This study applied Fisher Linear Discriminant for the first time as a linear discriminant in credit card fraud detection problem. Fisher Linear Discriminant or linear classifier (Christopher, 2006; Fisher, 1936; Fukunaga, 1990; McLachlan, 2004) utilizes dimension reduction method to find the best (D-1)-dimensional hyperplane(s) which can divide a D-dimensional space into two or more subspaces. It is a classic and popular supervised learning method which is commonly used in Face Recognition, Speech/Music Recognition, and Feature Extraction with some modifications (Alexandre-Cortizo, Rosa-Zurera, & Lopez-Ferreras, 2005; Liu & Wechsler, 2002; Witten & Tibshirani, 2011).

The main contributions of this study are introduction of Fisher Discriminant Function for the first time in credit card fraud detection literature and making a simple but effective modification to it to make it an empowered profit-driven classifier in this domain.

The outline of the rest of the paper is as follows: Section 2 reviews related works with detail, Section 3 introduce the methodology of Fisher Discriminant Analysis and improvement carried out in order to make it sensitive to individual profits. Section 4 illustrates the results of implementing the mentioned methods, whereas Section 5 concludes the paper and provides some possible future studies.

2. Related work

Since in this study our problem setup is built as developing a classifier which will help the business users to maximize their profit, here in this section instead of a thorough review of credit

card fraud or Fisher Discriminant Analysis publications, we focus on the rather narrow literature on cost sensitive or profit based learning. There is a little number of studies with regard to maximizing total profit (example-dependent) in implementing a classification tool, because as Elkan (2001) mentioned this kind of investigation is in its first steps.

An approach to take cost-sensitivity into account in building up a classifier is to adjust a threshold to make incorrectly classification of instance with higher cost of misclassification harder. In credit card fraud data set, since misclassification cost of fraudulent transactions as legitimate is much higher than misclassification cost of legitimate ones as fraudulent, there should be some modifications in cost matrix to perform better in minimizing total misclassification cost (Sheng & Ling, 2006; Zhou & Liu, 2006; Langford & Beygelzimer, 2005; Maloof, 2003). In real life problems like credit card fraud detection problem misclassification cost of instances may differ based on their classes. So in the mentioned studies, the authors developed a cost matrix showing classification cost of instances from class i as class j as $C(i,j)$. They showed that defining an appropriate cost matrix makes the learning models bias toward the instances with high misclassification cost. Maloof (2003) also indicated that adjusting a cost matrix have as same effect as sampling.

Another way of developing cost sensitive learning method is proposing a new model which is more sensitive to the important instances. Drummond and Holte (2000) developed a new decision tree which applies modified splitting criteria and pruning methods in order to sensitively classify instances with high cost of misclassification. In a similar study, Sahin et al. (2013) proposed a new cost sensitive decision tree which minimizes the misclassification cost while selecting the splitting attribute.

Another method to deal with cost-sensitive problems is using meta-heuristic algorithms with a fitness function taking into account the variable misclassification costs or profits. In a pioneer study, Duman and Ozcelik (2011) combined two well-known meta-heuristic algorithms – Genetic Algorithm (GA) and Scatter Search (SS) – called GASS. The proposed method could improve the performance of classification about 200% in terms of cost. In this study, the authors took the individually variable misclassification costs based on available usable limits.

As a purely relevant study, Duman and Elikucuk (2013a) applied migrating birds optimization (MBO) technique for first time in credit card fraud detection problem with the objective of maximizing total profit obtained by classifying the transactions instead of maximizing classification accuracy. The results show that the MBO algorithm has high performance in classifying most profitable transactions in comparison with the hybrid of Genetic Algorithm and Scatter Search (GASS). The authors on another research (Duman & Elikucuk, 2013b) proposed some modifications on neighborhood sharing function and benefit mechanism by which the total profit obtained could increase up to 94.2%. These results are based on real life data. The authors mentioned MBO as powerful meta-heuristic algorithm in credit card fraud detection problems.

3. Methodology

Below first Fisher Discriminant Analysis (FDA) and then the modification made on it are described.

3.1. Fisher Discriminant Analysis

Linear Discriminant Analysis (LDA) is a kind of supervised learning method by which the input region is divided into decision regions whose boundaries are called decision surfaces or decision boundaries. These decision boundaries are linear function of input

Download English Version:

<https://daneshyari.com/en/article/382849>

Download Persian Version:

<https://daneshyari.com/article/382849>

[Daneshyari.com](https://daneshyari.com)