Journal of Computational Science 27 (2018) 247–254

Contents lists available at ScienceDirect

Journal of Computational Science

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

Cost-sensitive Risk Induced Bayesian Inference Bagging (RIBIB) for credit card fraud detection

Akila S[*], Srinivasulu Reddy U

Department of Computer Applications, National Institute of Technology, Trichy 620015, India

ARTICLE INFO

Article history:
Received 20 October 2017
Received in revised form 26 April 2018
Accepted 18 June 2018

Keywords:
Credit card fraud detection
Bagging
Risk-based modeling
Bayesian inference
Weighted voting

ABSTRACT

Credit card fraud represents one of the biggest threats for organizations due to the probability of huge losses associated with them. This paper presents a cost-sensitive Risk Induced Bayesian Inference Bagging model, RIBIB, for credit card fraud detection. RIBIB proposes a novel bagging architecture incorporating a constrained bag creation method, a Risk Induced Bayesian Inference method as a base learner and a cost-sensitive weighted voting combiner. Experiments on Brazilian Bank data indicate 1.04–1.5 times reduced cost. Experiments on UCSD-FICO data exhibit robustness of the model in handling unseen data without any need for domain specific parameter fine-tuning.

1. Introduction

Increase in cashless transactions, especially card-based transactions are attributed to the improved availability of technology and the increased interest due to ease of use. Such increased adoption levels have made this domain lucrative and one of the major domains under the radar of fraudsters launching frequent exploitations or attacks. According to the Nilson Report (April 2017) [1], there are 293 million cardholders in the US, while 643.4 million cards are in circulation, resulting in an average of more than 2 cards per user. In the year 2016, it was observed that 34,260.6 million transactions have been performed, which brings the statistics to 66,089 transactions per second. Net loss due to fraudulent transactions amount to $2.17 billion in the same year. This clearly exhibits the elevated usage levels and escalated fraud levels. According to Unisys security index, credit and debit card frauds are the first concern for Americans, even superseding the fears of terrorism [2]. The reason ascribed to such a concern is the catastrophic impact of the worst-case scenarios and the notion that regardless of how low the occurrence of fraud may be, no one wants to find their hard-earned money tapped away by criminals.

The extensive adoption of cashless transactions has led to increased transaction data being generated, necessitating the use of automated machine learning models for identifying frauds [3]. Fraud detection is usually a supervised learning task, performed by classifiers. Prediction efficiency of classifiers depend on the efficacy of the data it is being trained upon. The huge transaction data generated due to customer purchases form the training data for classifiers. This huge amount of data forms a large training base so that the classifier can be effectively trained. This seemingly win-win situation also has several challenges to be tackled in-order to obtain a reliable classifier for usage in real-time. Intrinsic data imbalance contained in the transaction data poses a major challenge to the prediction process [4,5]. Although credit card transactional data is huge in number, the ratio of legitimate transactions to fraudulent transactions is very high, meaning that there exist too many legitimate transactions (called majority class) and comparatively too few fraudulent transactions (called minority class). This creates a bias during the classifier training process, leading to unreliable predictions [6,7]. Other issues affecting classifier performances include noise and concept drift [7,5]. The very nature of the domain tends to create several noisy entries in the data. Data elimination (undersampling) is usually the choice for eliminating data imbalance and noise [8,9]. However, in this domain, noisy data is not an error, instead, it is a behavioral change in the customer's spending pattern that needs to be recorded and utilized appropriately. Hence techniques to handle such data are to be explored, rather than data elimination. Concept drift is one of the intrinsic properties of domains dealing with data generated by customers [10]. Human behavior is ever changing, which tends to reflect in their spending patterns. This might be due to a change in their economic status or even due to inflation. However, the behavioral changes are never considered as factors when developing a fraud detection model.

[*] Corresponding author.
E-mail addresses: akila29@gmail.com (S. Akila), usreddy@nitt.edu (U. Srinivasulu Reddy).

https://doi.org/10.1016/j.jocs.2018.06.009
1877-7503/© 2018 Elsevier B.V. All rights reserved.

Though such models work well initially, in due course of time, error seeps in making the model obsolete. This reduces the dependability levels of the model, as the performance of the classifier model slowly deteriorates over time, finally becoming obsolete.

Performance of the classifiers are measured using several standard metrics like Accuracy, Precision, Recall, ROC, MCC, etc. However, analysis reveals that not all metrics reflects the performance of classifiers effectively when dealing with imbalanced data. Further, credit card fraud detection, being a business problem, has varied requirements. Business requirements are usually cost-based and such requirements exhibit considerable variation levels with the conventional metrics [11]. Several prediction models, even the ones that exhibit effective performance metrics are not adopted in real-time due to their low correspondence with business requirements. Business models usually necessitate predictions that incur low costs [12]. This cannot be directly mapped to the conventional classifier performance metrics, hence in this work, the cost is used as a metric and is considered as a major performance indicator during the prediction process.

This paper proposes a credit card fraud detection model named RIBIB, focused towards business goals and developed with high significance towards achieving low cost. The proposed credit card fraud detection model extends bagged ensembles by incorporating major enhancements in the process of bag creation, base learner creation, and result aggregation. The proposed RIBIB is incorporated with a modified bag creation approach to counter imbalance contained in the transaction data. It utilizes a Bayesian inference base learner, modified by incorporating the risk factors to provide risk-based results. Hence the proposed RIBIB model predicts based on the risk level of the transaction rather than the prediction probabilities. Further, the combiner model utilizes a cost-sensitive weighted voting mechanism for identifying the predictions with low-cost requirements. The remainder of this paper is structured as follows; Section 2 presents the related works describing prior works in credit card fraud detection domain, Section 3 presents an elaborate view of the proposed RIBIB model, Section 4 presents the results and discussions and Section 5 concludes the study.

2. Related works

Credit card fraud detection, even though a legacy problem, has evolved considerably so as to continually pose new challenges to the current research community. This section discusses recent works in the field of credit card fraud detection, categorized in terms of individual algorithm based detection models, stochastic meta-heuristic algorithm based detection models, ensemble-based detection models and cost and risk-based detection models.

A peer group based unsupervised fraud detection model that classifies fraudulent transactions by identifying diversity levels between previously similar transactions was proposed by Bolton et al. [13]. The downside of this model is that it uses a single attribute as the measurement parameter. Although additional attributes can be incorporated for analysis, the model is prone to the curse of dimensionality. A pattern identification based credit card fraud detection model using Apriori algorithm, called FraudMiner was proposed by Seeja et al. in [14]. This method identifies frequent item-set in the data to build a pattern repository, which contains signatures pertaining to fraud and legitimate transactions. An extension of this model, the Enhanced FraudMiner was proposed by Hegazy et al. in [15]. It uses LINGO, a clustering based algorithm for identifying the frequent patterns instead of Apriori. The major shortcomings of such pattern-based detection models is their inability to handle concept drift. A Cost-Sensitive Neural Network (CSNN) based bank fraud detection model was proposed by Ghobadi et al. in [16]. This model creates multiple neural network systems with varied topologies and identifies the best predictor from among them. The major disadvantage of this model is that it is computationally complex and is not robust. Neural Networks has been one of the most used models for credit card fraud detection. A combination model proposed by Brause et al. in [17], a fraud detection model for Mellon Bank proposed by Ghosh et al. [18], a GUI based fraud detection model, CARDWATCH, proposed by Aleskerov et al. [19], a combination model using Artificial Neural Networks and Bayesian belief networks proposed by Maes et al. [20] and fraud detection model by Dorronsoro et al. [21] are some of the most prominent models based on Artificial Neural Networks. Although neural network based models were found to be effective in detecting frauds, they are based on reducing false alarm levels, rather than cost and are also computationally intensive. This proves to be a major drawback, especially with the huge increase in the amount of data being generated by credit card transactions.

Increase in the amount of transaction data has resulted in several individual algorithm based models becoming obsolete as they necessitate high computational requirements. Meta-heuristic algorithm based models have become the most preferred choice due to their low computational requirements, faster processing capabilities, and stochastic operational process. An Artificial Immune System (AIS) based credit card fraud detection system called AIRS was proposed by Gadi et al. [22]. This is a cost-based fraud detection technique that uses Genetic Algorithm (GA) as the classifier. The model is entirely based on parameter fine-tuning to enhance the performance levels. This technique was observed to exhibit efficient performance levels on Brazilian bank data. However, analysis in terms of data imbalance is absent. An extension of the AIRS model [22], AFDM was proposed by Halvaiee et al. [23]. This model provides enhancement by improving several individual components of AIRS model. It improves the memory cell generation component, distance function, and the update component, and considers the properties of the datasets for evaluation. This model was observed to exhibit high performance, with improved fraud detection rates and lowered cost, however, it does so with the overhead of increased computations.

The use of Ensembles for fraud detection has currently gained importance due to the improved parallelization capabilities of the current computational systems. Some of the recent contributions in fraud detection using ensembles are listed below. A decision tree based bagged ensemble model for fraud detection was proposed by Zareapoor et al. [24], which yields a fraud catching rate of ~0.9 and a false alarm rate of <0.02. An ensemble created using a collection of Minimal Learning Machines (MLM) was proposed by Mesquita et al. [25]. This technique uses MLM as the base learner and also proposes two variants of MLM ensemble by incorporating Nearest Neighbour algorithm and Cubic Equation. The major advantage of this work is that it enables both classification and regression.

Cost-based models are currently on the rise due to the increased necessity towards creating models that align effectively with business goals. However, cost-based predictions have been analyzed as a viable option already been in use since late 90's. Cost based predictions are not used in isolation, instead, they are integrated with any of the previously discussed models to improve the model's performance. Works by Chan et al. [26,27] are some of the earliest works that deals with cost-based credit card fraud detection on skewed data. The meta-learning strategy proposed by Chan et al. involves a primitive data balancing strategy and cost based predictions. The limitation of this model is that it requires an initial analysis to determine the level of balancing and it involves multiple heterogeneous classifiers, making the model computationally intensive. A feature generation technique used for generating customer's spending patterns was presented by Bahnsen et al. in [28]. This model deals with completing the customer's spending profile by creating additional features for effective fraud detection. Time-frame based analy-