Credit Card Fraud Detection Using Convolutional Neural Networks

Kang Fu, Dawei Cheng, Yi Tu, and Liqing Zhang(✉)

Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
{fukang1993,dawei.cheng,tuyi1991,lqzhang}@sjtu.edu.cn

Abstract. Credit card is becoming more and more popular in financial transactions, at the same time frauds are also increasing. Conventional methods use rule-based expert systems to detect fraud behaviors, neglecting diverse situations, extreme imbalance of positive and negative samples. In this paper, we propose a CNN-based fraud detection framework, to capture the intrinsic patterns of fraud behaviors learned from labeled data. Abundant transaction data is represented by a feature matrix, on which a convolutional neural network is applied to identify a set of latent patterns for each sample. Experiments on real-world massive transactions of a major commercial bank demonstrate its superior performance compared with some state-of-the-art methods.

Keywords: Credit card fraud · Convolutional neural network · Imbalanced data

1 Introduction

With the rapid development of economy globalization in recent decades, credit cards are much more popular in commercial transactions. The corresponding problem of the credit card fraud emerges accordingly. Machine learning approaches have been proposed to overcome these challenges. Kokkinaki [4] proposed the decision tree and boolean logic functions to characterize the normal transaction modes so as to detect fraudulent transactions. However, some of the fraudulent transactions similar to the legitimate trading patterns can not be identified. So neural networks and Bayesian networks have been employed. Ghosh [2] used a neural network to detect credit card frauds. Bayesian belief networks and artificial neural networks have been also introduced to tackle the problem [6]. These models have been criticized for being overly complex to detect frauds and there has been a high probability of being over-fitting. In order to reveal the latent patterns of fraudulent transactions and avoid the model overfitting, we use a convolutional neural network to reduce the feature redundancy effectively.

How to generate features of credit card transactions successfully is one of the major challenges to machine learning approaches. Some aggregation strategies

A. Hirose et al. (Eds.): ICONIP 2016, Part III, LNCS 9949, pp. 483–490, 2016.
DOI: 10.1007/978-3-319-46675-0_53

[1,3,7,8] have been proposed to obtain the customer spending modes in the recent transactions. But these models can not describe the complicated patterns of consumer spending. Therefore, we design a novel trading feature called trading entropy based on the latest consumption preferences for each customer. In order to apply a convolutional neural network(CNN) to credit card fraud detection, we need to transform features into a feature matrix to fit the CNN model.

Besides, extremely imbalanced data is another issue in fraud detection. The random undersampling method for dominated class is a common technique to adjust the ratio of the minority. Unfortunately, it will inevitably dismiss valuable information. In this paper we generate synthetic fraudulent samples from real frauds by a cost based sampling method. Thus, we can get a comparable number of frauds with legitimate transactions for training.

In brief, the main contributions to this paper can be summarized as bellows:

1. We propose a CNN-based framework of mining latent fraud patterns in credit card transactions.
2. We transform transaction data into a feature matrix for each record, by which the inherent relations and interactions in time series can be revealed for the CNN model.
3. By combining the cost based sampling method in characteristic space, the extremely imbalanced sample sets are alleviated, yielding a superior performance of fraud detection.
4. A novel trading feature called trading entropy is proposed to identify more complex fraud patterns.

2 Methodology

In this section, we firstly provide a description of our CNN-based fraud detection framework. Secondly, we propose a novel trading feature. Thirdly, the cost based sampling method is elaborated. At the end of this section, the CNN model is employed to the problem of credit card fraud detection.

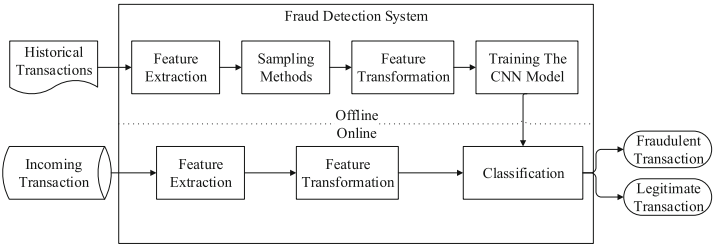

Fig. 1. The illustration of credit card fraud detection system.

2.1 Fraud Detection Framework

Our fraud detection framework is shown as Fig. 1, it consists of training and prediction parts. The training part mainly includes four modules: feature engineering, sampling methods, feature transformation and a CNN-based training procedure. The training part is offline and the prediction part is online. When a transaction comes, the prediction part can judge whether it is fraudulent or legitimate immediately. The detection procedure consists of feature extraction, feature transformation and the classification module.

For feature extraction, we adopt the aggregation strategies from [1,3,7,8]. In our system, we add trading entropy to the collection of traditional features so as to model more complicated consuming behaviors.

In the general process of data mining, we train the model after feature engineering. But a problem is that the data of credit card is extremely imbalanced. We propose a cost based sampling method to generate synthetic frauds.

Besides, in order to apply the CNN model to this problem, we need to transform features into a feature matrix to fit this model.

2.2 Feature Engineering

For traditional features, we can define the average amount of the transactions with the same customer during the past period of time as AvgAmountT. T means the time window length. For example, we can set T as different values: one day, two days, one week and one mouth, then four features of these time windows are generated. Table 1 shows the details of our feature types. Traditional feature types can not describe the complicated patterns of consumer spending. Therefore, we propose a new kind of feature called trading entropy. Assume in all transactions of the same customer during the past period of time before the current transaction, there are K kinds of merchant types, the total amount is TotalAmountT, the sum amount of the i-th merchant type is $AmountT_i (i = 1, 2, \ldots, K)$, the proportion of the i-th merchant type is p_i:

$$p_i = \frac{AmountT_i}{TotalAmountT} \tag{1}$$

The entropy of the i-th merchant type can be defined as $EntT$:

$$EntT = -\sum_{i}^{K} p_i log p_i \tag{2}$$

The above calculations only use previous transactions while the current transaction is not involved in. Then we add the current transaction to join the above calculation to obtain the current entropy: $NewEntT$. So the trading entropy is defined as $TradingEntropyT$:

$$TradingEntropyT = EntT - NewEntT \tag{3}$$

If the trading entropy is too large, it has a higher probability of being fraudulent.

2.3 Cost Based Sampling

The cost based sampling method is developed based on the following observation. The fraudulent transactions near the decision boundary have higher probabilities to generate more synthetic fraudulent samples. For the i-th fraud transaction, the number of frauds around i is defined as fd_i, and the number of normal transactions around i is defined as nd_i, the cost of the i-th transaction can be defined as $cost_i$. d_{ij} is the distance between the i-th fraud and j-th transaction. The number of neighborhoods of the i-th fraud can be limited by a transaction function $f(x)$ and a cutoff value. $f(x) = 1$ if $x < 0$ and $f(x) = 0$ otherwise and C is the cutoff.

$$cost_i = \frac{\sum_{j \in legitimate} f(d_{ij} - C)}{\sum_{k \in fraud} f(d_{ik} - C)} \tag{4}$$

After obtaining the cost of each fraudulent transaction, we use k-means algorithm to divide the frauds into some clusters. If we want to generate a new fraud sample, we choose a fraud transaction x_1 as the seed in accordance with the cost. Then we choose another fraud transaction x_2 from the same cluster as x_1. The new synthetic fraud sample can be generated as $newFraud = \alpha \cdot x_1 + (1 - \alpha) \cdot x_2$ where α is randomly generated between 0 and 1.

Table 1. The description of different feature types. The first seven feature types are traditional and the last one is proposed by us.

Feature types	Description
AvgAmountT	Average amount of the transactions during the past period of time
TotalAmountT	Total amount of the transactions during the past period of time
BiasAmountT	The bias of the amount of this transaction and AvgAmountT
NumberT	Total number of the transactions during the past period of time
MostCountryT	The mostly used country during the past period of time
MostTerminalT	The mostly used terminal during the past period of time
MostMerchantT	The mostly used merchant type during the past period of time
TradingEntropyT	The trading entropy gain during the past period of time

2.4 CNN Modeling

In this paper we apply the convolutional neural network to detect the frauds of credit cards, because the CNN model is suitable for training a large size of data and it has the mechanism to avoid the model over-fitting. Convolutional neural networks have been successfully applied to some fields, such as image classification and speech signal processing. But not all kinds of data are suitable for the CNN model. The method of feature transformations is proposed to adapt the CNN model. The features of credit card transactions can be partitioned into several groups. And each group has different features by different time windows.

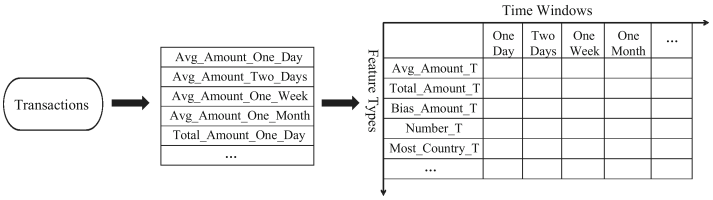

Fig. 2. The illustration of feature transformation. One dimensional features are generated from original attributes of transactions, then they are transformed into feature matrices.

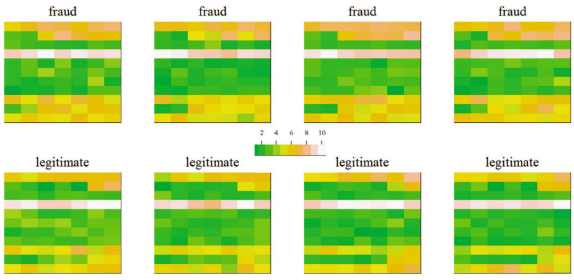

Fig. 3. The illustration of feature matrices. We randomly select four fraudulent and four legitimate samples to generate their heat maps after feature transformation.

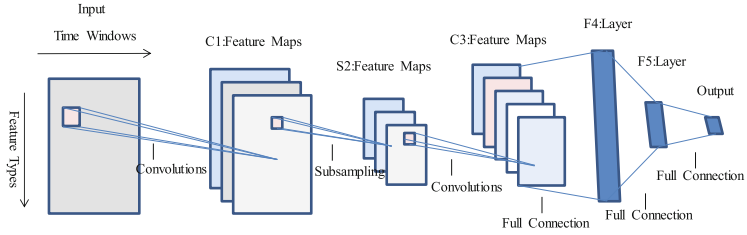

Fig. 4. The structure of our CNN model.

Two features of the same feature type by different time windows have strong relationship. Therefore, in the feature matrix, these two features are set in close positions. The original features are one dimensional. We can reshape them as a feature matrix where the rows have different feature types and the columns have different time windows. The procedure of transforming original features into a feature matrix is shown as Figs. 2 and 3.

These heat maps show the strong local correlation, in both row and column formats. According to the local correlation in the feature matrix, the CNN model can reduce the time complexity of data processing while retaining useful information. Our CNN structure is similar with the LeNet [5]. There are six

layers in total. The input is a feature matrix. And the first layer is a convolutional layer, the following is a sub sampling layer. The third layer is also a convolutional layer. And the last three layers are all full connection layers. Figure 4 shows the structure of our CNN model.

3 Experiments

This section is organized as three parts. Firstly, we introduce the dataset. Secondly, we show the importance of trading entropy. Finally, we demonstrate the best accuracy of the CNN model.

3.1 Dataset

To evaluate the proposed model, we use real credit card transaction data from a commercial bank. It contains over 260 million transactions of credit cards in a year. About four thousand transactions are labeled as frauds and the rest are legitimate transactions. The transaction data is divided into two sets. We take the data of the first 11 months as the training set and the data of the next month as the testing set. And we take the F1 score to evaluate the performance of models.

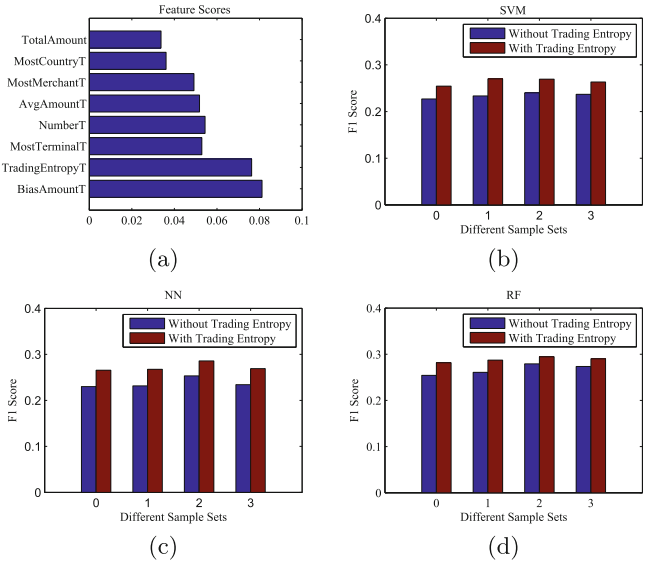

Fig. 5. Feature Evaluations. (a) Feature scores of different feature types. (b) The performances of SVM with trading entropy or not on different sample sets. (c) The performances of the neural network. (d) The performances of random forest.

3.2 Feature Evaluation

To evaluate the importance of trading entropy, we propose the feature score:

$$FeatureScore = \frac{1}{T}\sum_{t=1}^{T}\frac{\|u_t^f - u_t^l\|_1}{\sqrt{S_t^f + S_t^l}} \tag{5}$$

where T is the size of time windows, u_t^f and u_t^l represent the means of the fraud and legitimate samples for a given feature on t-th time window respectively, S_t^f and S_t^l are variances of the fraudulent and legitimate feature respectively. These scores are computed for each feature type. If the score is higher, this feature type is more important. Figure 5(a) shows the feature scores of different feature types. We can observe that the trading entropy ranks highly.

In order to demonstrate the efficiency of trading entropy better, we use different models and various sampling rates to obtain the performances with or without the trading entropy features. The testing results are illustrated in Fig. 5.

3.3 Model Evaluation

As shown above, the proposed trading entropy could significantly improve the classification accuracy. To relieve the problem of the imbalanced dataset, we employ the cost based sampling method to generate different number of frauds. In our experiments, we increase the fraudulent samples to 1, 2 and 3 times of the size of original frauds respectively. The samples of legitimate transactions are randomly undersampled. After feature engineering and data sampling, we evaluate the performance of the CNN model by comparing with other existing models. Figure 6 shows the comparison results. We can find that the cost based sampling method can make use of more legitimate data and alleviate the imbalanced problem. Besides, the CNN model on different sample sets achieves the best performance.

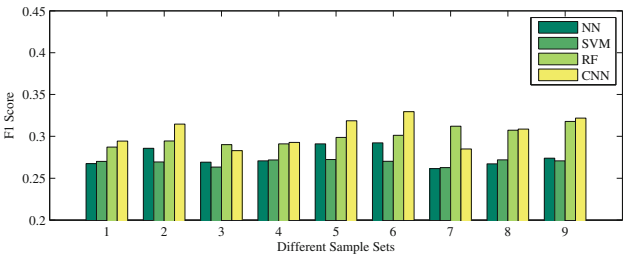

Fig. 6. The performances of different models on various sample sets.

4 Conclusion

In this paper, we introduce a CNN-based method of credit card fraud detection. And the trading entropy is proposed to model more complex consuming behaviors. Besides, we recombine the trading features to feature matrices and use them in a convolutional neural network. Experimental results from the real transaction data of a commercial bank show that our proposed method performs better than other state-of-art methods.

Acknowledgements. The work was supported by the National Natural Science Foundation of China (61272251), the Key Basic Research Program of Shanghai Municipality, China (15JC1400103) and the National Natural Science Foundation of China (91420302).

References

1. Bhattacharyya, S., Jha, S., Tharakunnel, K., Westland, J.C.: Data mining for credit card fraud: a comparative study. Decis. Support Syst. **50**(3), 602–613 (2011)
2. Ghosh, S., Reilly, D.L.: Credit card fraud detection with a neural-network. In: Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994, vol. 3, pp. 621–630. IEEE (1994)
3. Khandani, A.E., Kim, A.J., Lo, A.W.: Consumer credit-risk models via machine-learning algorithms. J. Banking Finan. **34**(11), 2767–2787 (2010)
4. Kokkinaki, A.I.: On a typical database transactions: identification of probable frauds using machine learning for user profiling. In: Proceedings of Knowledge and Data Engineering Exchange Workshop, 1997, pp. 107–113. IEEE (1997)
5. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
6. Maes, S., Tuyls, K., Vanschoenwinkel, B., Manderick, B.: Credit card fraud detection using bayesian and neural networks. In: Proceedings of the 1st International Naiso Congress on Neuro Fuzzy Technologies, pp. 261–270 (2002)
7. Ravisankar, P., Ravi, V., Rao, G.R., Bose, I.: Detection of financial statement fraud and feature selection using data mining techniques. Decis. Support Syst. **50**(2), 491–500 (2011)
8. Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B.: Apate: A novel approach for automated credit card transaction fraud detection using network-based extensions. Decis. Support Syst. **75**, 38–48 (2015)