Finance Fraud Detection With Neural Network

Yang YANG[1,a] Rong CHEN[2,b] Xiao BAI[3,c] DeHeng CHEN[4,d]
[1]King's own institute Sydney, Australia
[2]Central South University Changsha, China
[3]Efrei University Paris, France
[4]WuHan university Wuhan, China

Abstract—The payment card industry has grown increasingly with the development of online business. However, payment card fraud has become a serious problem around the world. Companies and banks lost huge amounts of dollars annually due to fraud. It is necessary to investigate a learning algorithm to detect fraud in finance transaction automatically. In this paper, we put forward a fraud detection algorithm by using neural network. The neural network model and final result will be described to show the superiority of this model.

1 INTRODUCTION

Finance Fraud can be defined as "intentional deception to secure unfair or unlawful gain". There are many ways to carry out a finance fraud due to the convenience of Internet. The most common fraud in daily life is credit card or debit card fraud. To address this problem, many researchers include data scientists and software engineers tried to develop a learning algorithm to find patterns from the fraud transaction and thus detect the potential frauds using this system. In past years, some machine learning algorithms is proposed to detect transaction frauds [1-3]. Machine learning algorithms is a process to build a mathematical and learning model based on training data to make predictions or decisions without being explicitly programmed. Some classical machine learning algorithms like logistic regression [6], Naïve Bayes [7] and support vector machine (SVM) [12] have been used in fraud detection task. However, these kinds of algorithms are unable to extract higher level of information from data. In this task, we built a deep neural network model which contain multiple perception layers to extract fraud information from data. The framework of neural network [4,5] and performance of it will be showed in the paper to inspire other scientists to have a further study.

A. Related Work

Many machine learning based algorithms have be proposed in recent years to detect frauds in transactions. In paper [1], the authors investigated the performance of naïve bayes, k-nearest neighbor and logistic regression on skewed credit card fraud data. In paper [2], Ong Shu Yee and Saravanan combined data mining technique and machine learning models to identify the genuine and non-genuine transactions from data. Sahil Dhankhad [3] and other contributors applied different machine learning algorithms in their study and employed a super classifier, which is an ensemble of several basic classifiers, to identify fraud in real-world dataset. However, as discussed before, most of these machine learning models are "shallow models" which unable to learn high level patterns iteratively from data. To handle this issue, we adopt deep neural network in this paper trying to extract high level patterns from real world data.

The proposed study of deep neural network on image classification has arose a revolution in many areas. Deep neural network has been applied to computer vision [9], natural language processing [10], speech recognition [11]. The success of deep neural network mainly results from three factors the advanced algorithms, the huge amounts of data and the rapid increasing computation resources like Gpu cards and Tpu Cards. Most of deep neural network methods has been achieve state of art performance.

B. Our Contribution

Considering the advantage of neural network, this paper proposes a deep neural network based algorithm to detect finance fraud. To have a better performance, certain data preprocessing methods like one hot encoding, standard normalization are used to train a better network. The following sections are organized as follows. Our Neural network models and parameter details are introduced in Section II. The third section presents discussed the dataset and preprocessing methods. In Section IV, we did experiments on proposed model and several classical machine learning models. Section V draws a conclusion of the whole paper.

[a]ab2613225@gmail.com [b]sianchan@163.com [c]xiao.bai@efrei.net [d]chendeheng611@gmail.com.

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

2 DEEP NEURAL NETWORK MODEL

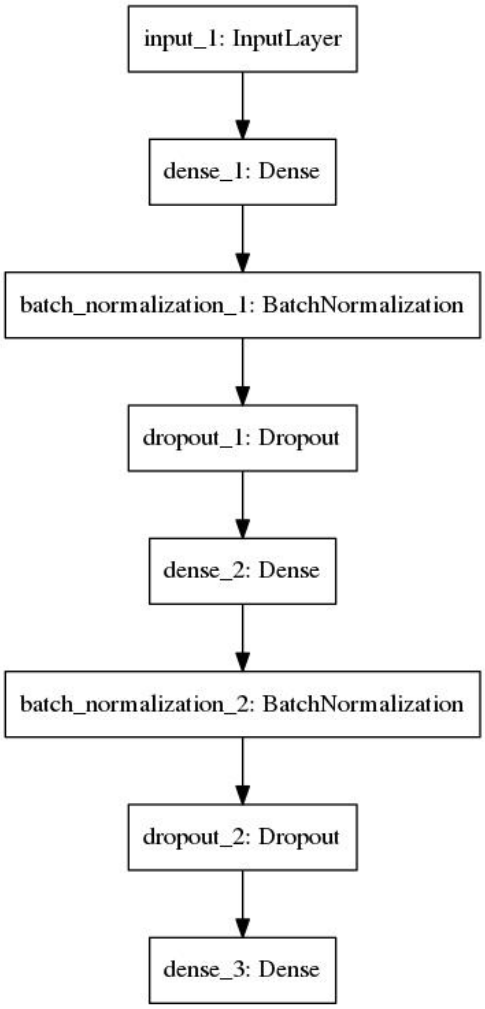

Figure 1. Our proposed model

The Figure 1 shows the proposed neural network model. The model mainly contains four layers, the first input layer, the second dense layer, the third dense layer and the last dense layer. The size of each layer is fixed and the concrete model parameters are listed in Table 1.

TABLE 1 Network parameters

Layers	**Layer Dimension**
Input layer	505
Dense_1 layer	512
Dense_2 layer	256
Dense_3 layer	1

The unit number of first layer is 505 which is equal to the dimension of preprocessed data. The last layer dimension is one because this task is a binary classification. The fraud transaction can be classified to be one and the normal transaction can be classified to be zero. The Dropout is regularization technique patented by Google in order to avoid overfitting. In this paper, two dropout layers with a rate of 0.3 are. added between layers. Batch normalization is a technique for improving the speed, performance, and the stability of artificial neural networks. It is used to normalize the layer by re-centering and re-scaling. The activation functions of this network for first three layers are Relu and the activation function of last layer is Sigmoid. Because the derivative of Relu will not attenuate when network go deeper and thus it is appropriate to use it in intermediate layers. The sigmoid is used because the task is a binary classification task.

3 DATA DESCRIPTION AND PREPROCESSING

In this section, the dataset and its preprocessing methods will be introduced. The dataset for our model is IEEE CIS [8] dataset, which is a famous dataset for transaction fraud detection. The dataset is mainly including two kinds of table: auxiliary transaction tables and identification tables. In this task, identification tables are not used because it has less attributes and it contributes little for our model. Hence, we will only introduce the transaction table.

A. Transaction table

As it is shown in Table 2, the transaction tables have 22 categorical features and 372 numeric features. The detail of these features is listed in table 1.

TABLE 2 TRANSACTION TABLE

Name	**Description**	**Type**
TransactionID	ID of transaction	ID
isFraud	binary target	categorical
TransactionDT	transaction date	time
TransactionAmt	transaction amount	numerical
card1-card6	card	categorical
addr1-addr2	address	categorical
M1-M9	anonymous features	categorical
P_email domain	purchaser email domain	categorical
R_email domain	receiver email domain	categorical
dist1-dist2	country distance	numerical
C1-C14	anonymous features	numerical
D1-D15	anonymous features	numerical
V1-V339	anonymous features	numerical

In preprocessing part, the first part is to do data cleaning. Some columns with majorities of Nan values should be removed and some columns with high correlated also should be removed. In missing value filling part, the Nan values should be replaced with some numeric values like 0. Otherwise, it is unable for computer to process this kind of data. Then for categorical features and numeric features, different processing method are used. For numeric data, it should be normalized. In this task, the standard normalization is used. Standard normalization can be done by subtracting the mean values of data and then divide by standard deviation.

For categorical data, it should be encoded by one hot encoding. One hot encoding is a group of bits among

which the legal combinations of values are only those with a single high (1) bit and all the others low (0).

4 EXPERIMENTS

The training of our proposed model can be done via CPUs or GPUs. However, it is 10 more times faster to using GPUs like Nvidia Rtx Titan for training model. In addition, the GPU accelerating packages like Cudnn and Cuda can be installed to quicken the speed of training phase and inference phase. To help train model faster, the parameters for training model should be set carefully. In Table 3, it lists the parameters for neural network. For example, the optimizer is Nadam, which is an adaptive optimizer so that the model will converge in a quick speed.

TABLE 3 Model parameters

Parameter	**Parameter Description**	**Value**
Optimizer	Optimizer of network	Nadam
learning_rate	learning rate	0.0001
epoch	iteration number	8
batch_size	Batch size	2048

TABLE 4 Performance of different models

Models	**Auc Roc Score**	**Accuracy**
Naïve Bayes	0.813	0.868
Logistic Regression	0.836	0.904
GBDT	0.852	0.935
Neural Network	**0.904**	**0.974**

From table 4, it is easy to find out that the Neural network based model outperforms the other three models on both auc roc score and accuracy score, which means the proposed can detect fraud more accurately than other three models.

The figure 2 shows the training loss curve via different epochs. It is obvious that our model can converge really quick in just 7 epochs.

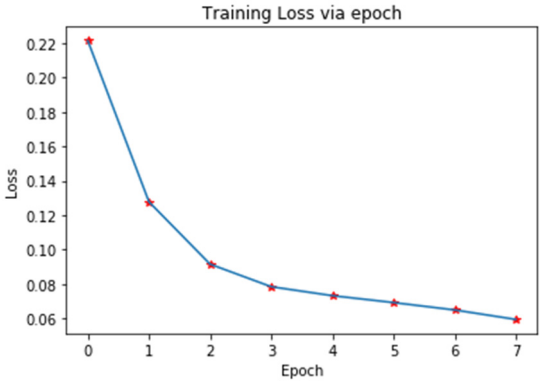

Figure 2. The training loss curve

5 CONCLUSIONS

This paper presents a finance fraud detection algorithm based on neural network. In first part, we introduce the related work in fraud detection studies and deep neural networks. In section II, the neural network model is displayed in Figure 1 and the function of each layer is described. In third section, the data description and the preprocessing techniques like one hot encoding, standard normalization are introduced. In experiment part, the deploy environment is proposed firstly. Then the model parameters and performance of different models are also shown to show the superiority of the proposed model.

ACKNOWLEDGEMENT

The author Yang YANG thanks Rong CHEN for help in feature engineering and data preprocessing. Then it is nice for Xiao BAI to tune the parameter of neural network models. Last but not least, we thank Deheng CHEN to carry out in Ubuntu environment and record some key results.

REFERENCES

1. Awoyemi, John O., Adebayo O. Adetunmbi, and Samuel A. Oluwadare. "Credit card fraud detection using machine learning techniques: A comparative analysis." 2017 International Conference on Computing Networking and Informatics (ICCNI). IEEE, 2017.
2. Yee, O. S., Sagadevan, S., & Malim, N. H. A. H. (2018). Credit card fraud detection using machine learning as data mining technique. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(1-4), 23-27.
3. Dhankhad, S., Mohammed, E., & Far, B. (2018, July). Supervised machine learning algorithms for credit card fraudulent transaction detection: a comparative study. In 2018 IEEE International Conference on Information Reuse and Integration (IRI) (pp. 122-125). IEEE.
4. Beale, H. D., Demuth, H. B., & Hagan, M. T. (1996). Neural network design. Pws, Boston.
5. Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. In Neural networks for perception (pp. 65-93). Academic Press.
6. Tolles, J., & Meurer, W. J. (2016). Logistic regression: relating patient characteristics to outcomes. Jama, 316(5), 533-534.
7. Murphy, K. P. (2006). Naive bayes classifiers. University of British Columbia, 18, 60.
8. https://www.kaggle.com/c/ieee-fraud-detection/data
9. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks.

In Advances in neural information processing systems (pp. 1097-1105).

10. Socher, R., Bengio, Y., & Manning, C. D. (2012, July). Deep learning for NLP (without magic). In Tutorial Abstracts of ACL 2012 (pp. 5-5). Association for Computational Linguistics.
11. Deng, L., & Platt, J. C. (2014). Ensemble deep learning for speech recognition. In Fifteenth Annual Conference of the International Speech Communication Association.