Semi-Supervised Classification on Credit Card Fraud Detection using AutoEncoders

by Nur Dzakiyullah

Submission date: 20-Jul-2022 09:48AM (UTC+0700)
Submission ID: 1872850659
File name: 16-72-7-PB_3.pdf (834.96K)
Word count: 3715
Character count: 20097

Journal of Applied Data Sciences
Vol. 2, No. 1, January 2021, pp. 01-07

Semi-Supervised Classification on Credit Card Fraud Detection using AutoEncoders

Nur Rachman Dzakiyullah [1,*], Andri Pramuntadi [2], Anni Karimatul Fauziyyah [3]

[1] Faculty Computer, Department of Informatics, Alma Ata University,Yogyakarta, Indonesia
[2] Faculty Computer, Department of Informatics, Alma Ata University,Yogyakarta, Indonesia
[3] Faculty Computer, Department of Informatics, Alma Ata University,Yogyakarta, Indonesia
[1] nurrachmandzakiyullah@almaata.ac.id*; [2] andripramuntadi@almaata.ac.id; [3] anni.karim@almaata.ac.id
* corresponding author

(Received December 21, 2020, Revised January 3, 2021, Accepted January 14, 2021, Available online January 15, 2021)

Abstract

The use of credit cards for online purchases has increased dramatically and led to an explosion in credit card fraud. Credit card companies need to be able to identify fraudulent credit card transactions so that customers are not charged for items they do not buy. In this study, we will use semi-supervised learning and combine it with AutoEncoders to identify fraudulent credit card transactions. In this paper, we will implement the use of T-SNE to visualize fraud and non-fraud transactions, then improve the visualization using autoencoders. Classification report proved that it is possible to achieve very acceptable precision which is almost 100% rate using semi-supervised classification to detect credit card fraud. This study is based on a dataset achieved by kaggle "Credit Card Fraud" and has been improved by using the ideal research model available.

Keywords: Data Science. Semi-Supervised Classification, Credit Card Fraud, AutoEncoders

1. Introduction

Fraud is a serious crime that typically involves money and a business transaction of some kind. People are less protective during this day in the era of e-commerce when engaged in any transaction. Recently, the use of credit cards for online purchases has increased, people are not generally aware of having fraud transactions. The security of a credit card relies on the physicality of the card and the privacy of the credit card number. Globalization and the growing use of the Internet for online commerce have led to a major expansion of worldwide credit card purchases. A rapid increase in the number of credit card purchases has also contributed to a major rise in fraudulent activities. Credit card fraud is a broad-ranging term for theft and fraud performed as a fraudulent funding source in a specific interaction using a credit card. Theft and fraud are committed in a given transaction using a payment card as a fraudulent source of funds. A vast range of methods to conduct theft are used by Credit Card Fraudsters. To successfully fight credit card fraud, it is important to first understand the processes for detecting credit card fraud. Due to numerous credit card fraud monitoring and avoidance mechanisms, credit card fraud has stabilized a lot over the years.

However, the use of false transactions to defraud bank cash by cardholders. External card fraud, on the other hand, is primarily expressed in the use of a stolen, fraudulent, or counterfeit credit card to consume or use cards to get cash in concealed ways, such as the purchasing of valuable, limited amounts of products or items that are easy to sell in cash. This article will specifically explore & analyze the development of credit card fraud detection using machine learning.

2. Literature Review

2.1. Credit Card Fraud

When using standard techniques, detecting credit card fraud is a difficult job to do so the creation of credit card fraud has been important, whether in the academic or business environment. Richard and David[1] indicated that two types of credit card fraud are available: application and behavioral fraud. Applicational fraud is where the suspect or we may call it fraudsters who receive new cards from issuing firms using fake details or information from other individuals. Behavioral fraud, but again, can again be categorized into 4 types: mail theft, stolen/lost card, fake card, and non-fraud card owners.

Mail theft scams occur when fraudsters intercept credit cards in the mail or steal personal information from banks and credit card statements before they reach the cardholder. [2] Stolen/lost card fraud occurs when a fraudster obtains a credit card or accesses a lost card through wallet/wallet theft. However, with the increasing use of online transactions, the number of counterfeit cardholders (not con artists) has increased rapidly. In both types of fraud, credit card details are obtained without the knowledge of the cardholder, and then create fake cards, or use the information to expose the cardholder to letterless transactions, phone calls, or cardholder internal information is obtained through a variety of methods. Various ways, such as employees stealing information via unauthorized wipers, phishing scams, or cardholders, do not prevent fraud, credit card details are used remotely to perform fraudulent transactions.

For ten years, credit card fraud has been growing[3]. The most popular form of credit card theft was stolen and counterfeit cards in the 1970s, in which a physical card was stolen and used. Later, in the 1980s and 1990s, mail order/phone order became widespread. Digital crime has recently migrated to the Internet, which can be perpetrated internationally, offering anonymity, coverage, and speed, and fraud. This is no longer the case for the use of technology by a single person, but with an increasing and coordinated group of actors actively evolving their own technology.

Botan and Hand[4] have written literature on the prevention of credit card fraud, which makes it impossible to share ideas and suppress possible fraud detection technologies. On the one hand, it is difficult for researchers to collect data sets on credit card transactions that obstruct transaction analysis, on the other hand, fraudsters are able to acquire information and avoid detection through detection methods that have not been commonly debated in society. There was some positive debate between Bolton[5] and Provost[6] about credit card fraud identification studies in 2002.

Credit card processing records typically contain properties of numbers and groups. A standard digital attribute, such as business code, business name, transaction date, and other type attributes, is the transaction number. Data sets of hundreds or thousands of categories may depend on any of these categorical variables. The use of numerous mathematical methods, machine learning and data processing has resulted in the combining of many broad numerical values and categorical attributes[7]. We are presented with the difficulty of using numbers and intelligent classification attributes in this report. By merging cardholder transaction details over a specific period of time, certain attributes are generated. The first is the inevitable disparity in the data set's allocation of divisions. The number of illegitimate transactions in the real world is (unfortunately) far lower than that of legal transactions, but their inaccurate distribution can cause model prediction bias [8]. Second, the vast amount of uncommunicated samples that have given rise to such exchange exclusion policies or because messages have failed to reach the banking system. Accordingly, the denied transactions cannot be linked to theft in this situation.

2.2. Machine Learning

Machine learning (ML) technology has been an alternative approach for creating automatic FDI in recent years. ML requires different means of theft to be considered. A fraud detection challenge lends itself to supervised classification tasks from a machine learning viewpoint, which is programmed to assess whether a new transaction is valid or fraudulent. In order to match the inductive class distribution model, the supervised

ML algorithm training process is expected. The general model to be designed, however, has three difficulties, which render the method a challenging task. Then several samples were found unimportant, likely because such trade policies were refused or because they failed to interact with the banking system. Accordingly, the denied transactions cannot be linked to theft in this situation. Finally, since a huge number of transactions and a large volume of data have to be handled on a regular basis, flexible training methods need to be used. With all this in mind, unlabelled samples that incorporate supervision and semi-supervision processes and often operate in large data environments should be able to use the learning approaches used to design automated FDS.

Many methods have been used to assess credit card theft in recent years. ANN or artificial neural networks, like many other applications, have been widely used in this respect. Centered on FDS based on MLP or multilayer perceptron, Eleskov et al.[9], Dorronsoro[10], Kim et al[11]. Olszewski[12] read the self-directed map later on and Ogwueleka[13] read about the use of two neural networks. Decision trees[14], vector support machines[15],[16], logistic regression[17] and Bayesian learning[18] are other methods that can be used. Any of the previous approaches, such as Whitrow et al.[19], were contrasted by other scholars. The authors merged data aggregation strategies with different classification algorithms, such as supporting vector machines, logistic regression, Bayesian learning, decision-making, and decision-making. Tree Committee (called random forest). Comparisons were made in the analysis of Bhattacharyya et al.[20] between vector support machines, logistic regression and random forests. In[21], "random forest", "logistic regression" and "artificial neural networks" were contrasted by the developers of the real-time credit card fraud process.

3. Method

Machine learning can be defined as the complex method of observing a data collection of the best and most important trends, relationships, or connections that can be used to predict results from unseen data. Broadly speaking, three separate mechanisms of machine learning exist:

- *Supervised Learning* is the training method on labeled data sets for a machine learning algorithm. Configured data with known target variables. The model aims at evaluating the relationship between the independent variable and the dependent variable in this methodology. Classification, regression, and inference provide examples of supervised learning.
- *Unsupervised Learning is* the method of training a machine learning algorithm on a data set of undefined target variables is unsupervised learning. The model attempts to identify the most important trends in data or data fragments in this methodology. Cluster segmentation, dimension elimination, etc are forms of unsupervised learning.
- *Semi-Supervised Learning* is a mixture of a supervised learning process and an unsupervised learning process, where the model is often trained with unlabelled results. The unsupervised learning attribute is used in this method to research the best possible representation of data, and the attribute of directed learning is used to study the relationships in that representation and then to make predictions.

When the data collection is very unbalanced, semi-supervised learning is extremely useful. The aim is to introduce a paradigm that suits the dominant class well into the labels. The model would not fit a tag that constitutes a scam, and we can set criteria on how the model has to be adapted or not fraudulent by adding a threshold. In this study, using the semi-supervised learning process, we will explain how to execute classification tasks. In order to analyze data representations, this approach utilizes an autoencoder and then trains a basic linear classifier to assign the data collection into its own class.

3.1. Visualize Fraud & Non-Fraud Transactions

We can use T-SNE to visualize the essence of fraudulent and non-fraudulent transactions. Random neighbor embedding T-SNE or t-distribution is a nonlinear dimensionality reduction technique, which is particularly useful for visualizing high-dimensional data sets. It is commonly used in the processing of images, NLP, genetic information, and voice processing. The discrepancy between two distributions is reduced by the T-

SNE: one is used to measure the pairwise similarity of the input object, the other is used to measure the pairwise similarity of the embed's corresponding low-dimensional point.

In this way, t-SNE maps to a lower-dimensional space the multi-dimensional data and tries to identify patterns in the details. Transactions reflect the following points. non-fraudulent transactions are represented in green, and red is shown for fraudulent transactions. The elements that T-SNE removes are these two axes. Many non-fraud transactions are very similar to fraudulent transactions, as can be seen from the figure below, making it impossible to classify correctly from the model.

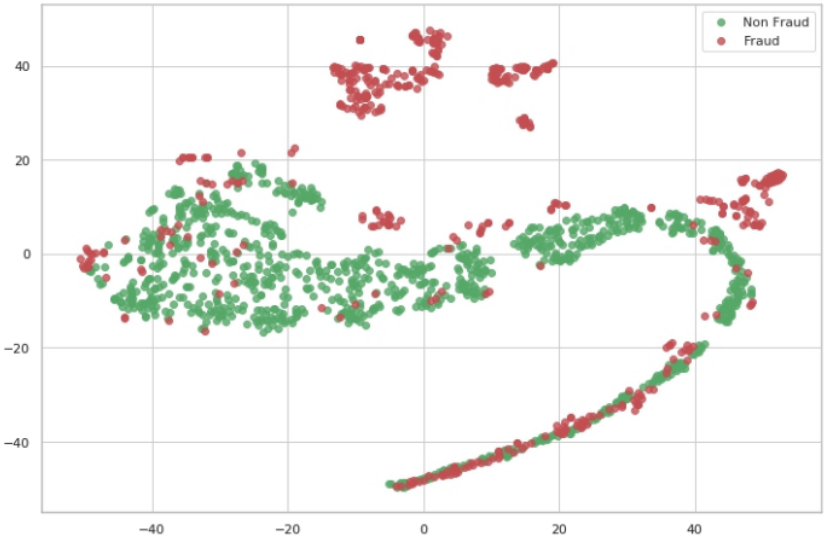

Figure. 1. T-SNE Actual Representation

3.2. AutoEncoders

Why can you need an auto-encoder? We need to remember, first of all, what an automated encoder is. A special neural network architecture whose output is the same as the input is the autoencoder. Autoencoders are qualified to learn low-level input data representations in an unsupervised manner. Fit this low-level attribute to the individual results, then. A regression task in which the network is expected to predict its input is an autoencoder (in other words, to model an identity function). In the centre, the network has tight resistance to multiple neurons, forcing them to construct powerful representations, compressing the data into low-dimensional data, and the decoder will replicate the original input using this low-dimensional code.

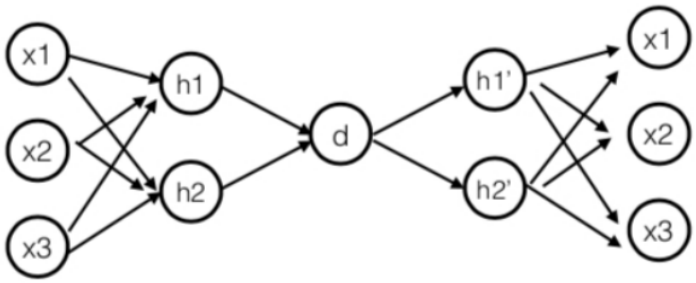

Figure. 2. AutoEncoders Model

We're going to create an auto-encoder model that demonstrates a non-fraud case model only. The model would strive to think about the correct interpretation of cases of non-fraud. To produce representations of fraud cases, the same model will be used and we assume the model to be distinct from cases of non-fraud. The benefit of this approach is that to learn a decent representation, we don't need too many data samples. To train an autoencoder, we can only use 2000 lines of non-fraud cases. Even, we don't have a long time to run this model. Increases after the model. For the study model, we are interested in obtaining a latent representation of the inputs. Using trained model weights, you can use them. We're going to build another sequential layer network, and we're just going to add weights that inhabit the third layer where a latent representation resides.

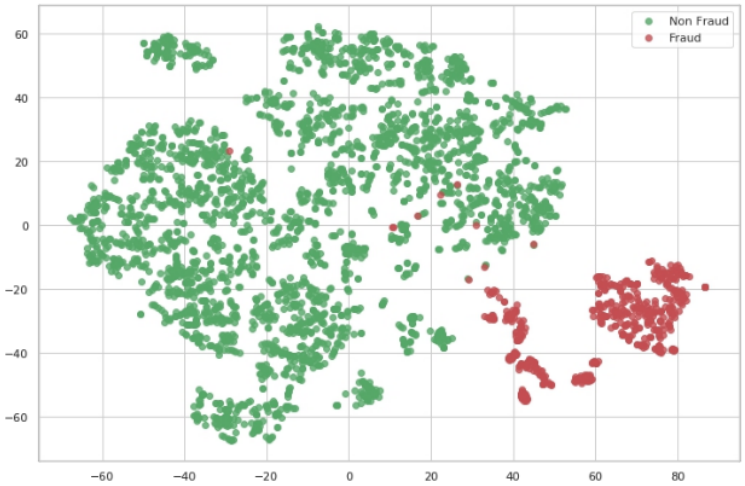

Figure. 3. T-SNE Latent Representation

3.3. Visualize the latent representation: Fraud & Non-Fraud

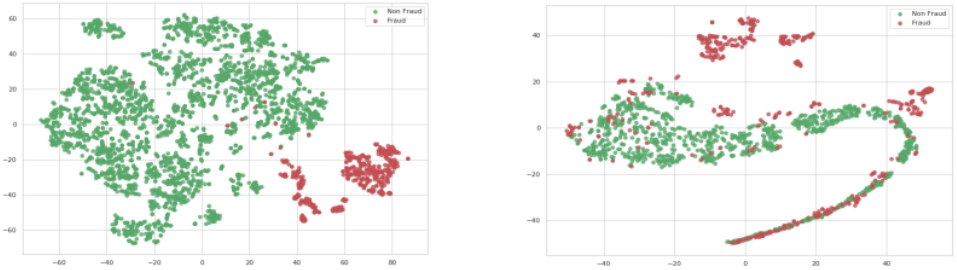

Figure. 4. T-SNE comparison between actual & latent representation

We can imagine the nature of the cases of fraud and non-fraud until a possible sign is received. Fraudulent transactions and non-fraudulent transactions, as can be seen from the picture above, are now very simple and can be distinguished linearly. Now in order to define this, we don't need a complex model and we can use even simpler models to make predictions. Situations before and after corrupt and non-fraudulent transactions are as follows. We just need to train a basic linear classifier on the data set after obtaining these results to get the accuracy, precision, f1 value, and recall score from the model. See table. 1 for classified reports.

Table. 1. Classification Report

	Precision	Recall	f1-Score	Support
0.0	0.98	1.00	0.99	754
1.0	1.00	0.87	0.93	119
Micro avg	0.98	0.98	0.98	873
Macro avg	0.99	0.94	0.96	873
Weighted avg	0.98	0.98	0.98	873

4. Results

For this study, it was proved that the semi-supervised learning model performed very well according to the result. The Classification achieves a very good precision on all labels correctly. But, still, there is much more possibility and improvement we can do about this study. The combination between T-SNE and AutoEncoders is just the beginning, we can use other Semi-supervised methods like Gaussian Mixture and Isolation Forest.

The results obtained are very supportive of the possibility that credit card fraud can be very easily classified. This will really help a lot of parties such as for credit card owners, they can easily detect and find out potentials prone to credit card fraud or for credit card providers they can take advantage of the results of this research to increase the security of their credit card provider. The results of this research are only the culmination of an iceberg, which means that there are still a lot of possibilities and of course in the future whether we can develop to overcome this type of fraud or even as the era of fraud like this develops, further research is needed.

5. Discussion and Conclusion

Future studies could investigate the idea that intelligent dependent attributes could be developed to help more reliably identify transactions. Based on previous research, we created derived attributes, but future work could help to carry out more extensive research on the attributes most suitable for fraud modeling, including problems with transaction aggregation[22]. How card fraud with multiple fraudulent transactions differs from cards with a small number of fraudulent transactions is another problem that needs to be studied. Our data limitation, as stated above, is that it is not possible to use the exact time stamp data after the credit card transaction date [23]. Before the credit card is withdrawn, future research may look at differences in the sequence of fraudulent and legal transactions. In future studies, differences in fraud between different types of fraud can also be examined, such as differences in behavior between stolen cards and counterfeit cards.

Another matter is alternative methods of dividing data into training and testing. A random sampling of data in training and testing, as used in this study, assumes that the cheating pattern will not change significantly over the expected pattern of implementation. Since fraudsters' mechanisms are becoming increasingly complex and these mechanisms can be changed from time to time to avoid detection, the assumption of a stable timeout mode may not apply. Therefore, data bias may be important to consider. Training and testing data can be prepared to test the predictive ability of the trained model in a later phase to better fit the way the developed model is used in re-application. It is useful to examine the extent of the deviation of the concept and whether the patterns of fraud persist over time, with the availability of data covering a long period.

References

[1] E. Aleskerov, B. Freisleben, B. RaCARD WATCHTCH: a neural work-based database mining system for credit card fraud detection, in computational intelligence for financial engineering, Proceedings of the IEEE/IAFE, IEEE, Piscataway, NJ, 1998, pp. 220–226

[2] CapitalOne Identity theft guide for victims, RetrievedJanuary 10, 2009, from http://www.capitalone.com/fraud/IDTheftPackageV012172004We.pdf?linkid=WWW_Z_Z_Z_FRD_D1_01_T_FIDTP

[3] K. Williams, The Evolution of Credit Card Fraud: Staying Ahead of the Curve,eFunds Corporation, 2007.

[4] R.J. Bolton, D.J. Hand, Unsupervised profiling methods for fraud detecti on, Conference on Credit Scoring and Credit Control, Edinburgh, 2001.

[5] R.J. Bolton, D.J. Hand, Statistical fraud detection: a review, Statistical Science 17(3) (2002) 235–249.

[6] F. Provost, Comment on Bolton and Hand, Statistical Science 17 (2002) 249–251.

[7] C. Whitrow, D.J. Hand, P. Juszczak, D. Weston, N.M. Adams, Transaction Aggregation as a strategy for credit card fraud detection, Data Mining and Knowledge Discovery 18 (1) (2009) 30–55.

[8] N. Chawla and K. Bowyer, "SMOTE: synthetic minority oversampling technique,"arXiv preprint arXiv: . . ., vol. 16, pp. 321–357, 2011.

[9] E. Aleskerov, B. Freisleben, and B. Rao, "Card watch: A NeuralNetwork Based Database stern for Credit Card Fraud Detection," Proceedings of the IEEE/IAFE, pp. 220–226, 1997

[10] J. R. Dorronsoro, F. Ginel, C. S anchez, and C. Santa Cruz, "Neural Fraud detection in credit card operations," IEEE Transactions on neural networks, vol. 8, no. 4, pp. 827–834, 1997

[11] M.-J. Kim and T.-S. Kim, "A neural classifier with fraud density map for effective credit card fraud detection," Intelligent Data Engineering and Automated LearningIDEAL 2002, pp. 21–30.

[12] D. Olszewski, "Fraud detection using self-organizing map visualizing the user profiles," Knowledge-Based Systems, vol. 70, pp. 324–334,2014.

[13] F. Ogwueleka, "Data mining application in the credit-card Fraud detection system," Journal of Engineering Science and Technology, vol. 6, no. 3, pp. 311–322, 2011.

[14] Y. Sahin, S. Bulkan, and E. Duman, "A cost-sensitive decision tree approach for fraud detection," Expert Systems with Applications, vol. 40, no. 15, pp. 5916–5923, 2013

[15] Qibei Lu; Chunhua Ju, "Research on Credit Card Fraud DetectionModel Based on Class Weighted Support Vector Machine," Journal Of Convergence Information Technology, vol. 6, no. 1, pp. 62–68, 2011

[16] Dheepa and R. Dhanapal, "Behavior-BasedCredit Card FraudDetection Using Support Vector Machines," ICTACT Journal on SoftComputing, vol. 6956, no. July, pp. 391–397, 2012.

[17] S. Jha, M. Guillen, and J. Christopher Westland, "Employing transaction aggregation strategy to detect credit card fraud," Expert Systems withApplications, vol. 39, no. 16, pp. 12 650–12 657, 2012.

[18] S. Panigrahi, A. Kundu, S. Sural, and a. K. Majumdar, "Credit card fraud detection: A fusion approach using Dempster-Shafer theory and bayesian learning," Information Fusion, vol. 10, no. 4, pp. 354–363,2009.

[19] C. Whitrow, D. J. Hand, P. Juszczak, D. Weston, and N. M. Adams, "Transaction aggregation as a strategy for credit card fraud detection," Data Mining and Knowledge Discovery, vol. 18, no. 1, pp. 30–55, 2009.

[20] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data Mining for credit card fraud: A comparative study," Decision support systems, vol. 50, no. 3, pp. 602–613, 2011.

[21] V. Van Vlasselaer, C. Bravo, O. Caelen, T. Eliassi-Rad, L. Akoglu, M. Snoeck, and B. Baesens, "APATE: A Novel Approach for Auto-mated Credit Card Transaction Fraud Detection using Network-BasedExtensions," Decision Support Systems, vol. 75, pp. 38–48, 2015.

[22] C. Whitrow, D.J. Hand, P. Juszczak, D. Weston, N.M. Adams, Transaction Aggregation as a strategy for credit card fraud detection, Data Mining and Knowledge Discovery 18 (1) (2009) 30–55.

[23] M. Imron and S. A. Kusumah, "Application of Data Mining Classification Method for Student Graduation Prediction Using K-Nearest Neighbor (K-NN) Algorithm," IJIIS Int. J. Informatics Inf. Syst., vol. 1, no. 1, pp. 1–8, 2018, doi: 10.47738/ijiis.v1i1.17.

Semi-Supervised Classification on Credit Card Fraud Detection using AutoEncoders

ORIGINALITY REPORT

3%	2%	2%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	José R. Dorronsoro. "Architecture Selection in NLDA Networks", Lecture Notes in Computer Science, 2001 Publication	<1%
2	ep.liu.se Internet Source	<1%
3	towardsdatascience.com Internet Source	<1%
4	"World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada", Springer Science and Business Media LLC, 2015 Publication	<1%
5	Anisah H. Nizar, Zhao Yang Dong, Pei Zhang. "Detection rules for Non Technical Losses analysis in power utilities", 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 Publication	<1%

6	e-journal.unipma.ac.id Internet Source	<1%
7	galaxydatatech.com Internet Source	<1%
8	www.coursehero.com Internet Source	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		