DOI: 10.12731/2070-7568-2022-11-3-17-28
UDC 330.4

DETECTION OF CREDIT CARD FRAUDS WITH MACHINE LEARNING SOLUTIONS: AN EXPERIMENTAL APPROACH

C. Mabani, A.A. Tuskov, E.V. Shchanina

Purpose *– propose an experimental way to create ML solutions to the problem of detecting credit card fraud.*

Method or methodology of the work: *the article uses machine learning (ML) and data mining methods*

Results: *the paper showed that machine learning (ML) and data mining techniques are effective in improving fraud detection accuracy. The study proposes an experimental way to create ML solutions to the problem aimed at minimizing financial losses by monitoring the client's behavior when using credit cards. The model is tested on a publicly available dataset available to the research community in terms of detection accuracy.*

The sphere of application of the results: *in practice, it is advisable to use the results when planning effective strategies for detecting fraud in credit cards.*

Keywords: *fraud; credit cards; machine learning; experimental approach*

ОБНАРУЖЕНИЕ МОШЕННИЧЕСТВА С КРЕДИТНЫМИ КАРТАМИ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ: ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД

К. Мабани, А.А. Тусков, Е.В. Щанина

Цель *– предложить экспериментальный способ создания ML-решений проблемы обнаружения мошенничества с кредитными картами.*

Метод или методология проведения работы: *в статье использованы методы машинного обучения (ML) и интеллектуального анализа данных.*

Результаты: *в статье было показано, что методы машинного обучения (ML) и интеллектуального анализа данных эффективны в повышении точности обнаружения мошенничества. В исследовании предлагается экспериментальный способ создания ML-решений проблемы, направленных на минимизацию финансовых потерь путем мониторинга поведения клиента при использовании кредитных карт. Модель тестируется на общедоступном наборе данных, доступном для исследовательского сообщества с точки зрения точности обнаружения.*

Область применения результатов: *на практике полученные результаты целесообразно применять при планировании эффективных стратегий выявления мошенничества в кредитных картах.*

Ключевые слова: *мошенничество; кредитные карты; машинное обучение; экспериментальный подход*

Introduction

The global rate of online fraudulent transactions has grown to 1% according to the UK finance report [26]. In the United States, fraudulent claims only in healthcare and insurance led to financial losses amounting to 98 billion and 300 billion a year, respectively. Machine learning (ML) and Data Mining techniques were shown to be efficient to improve the detection of credit card frauds, see e.g. [17, 5, 7].

An example of a real data set of credit card transactions made by the EU cardholders in 2013 during 2 days includes 284,807 transactions in which 492 or 0.17% transactions were recorded as fraud. This data set has been made available for research. The data are represented by 31 variables made of principal components to make the data anonymous. The variables Time, Amount and Class are made available. The Class is the targeted variable, where 1 is fraud and 0 normal transaction. The data set is unbalanced and so requires a special technique estimating the detection accuracy, sensitivity and specificity [12, 2, 18]. The research on such a data set is limited because of using the principal components which cannot be directly explained, as described in [13].

First we analyse Random Forest (RF) models well-known in the literature for provision of efficient solutions to real-world problems including credit card fraud detection. In particular we analyse (i) how samples of customer's payment transactions impact the detection accuracy and (ii) how a fraudulent transaction can be detected most reliably. According to [2] the random forests perform better than other models such as ANN which can be affected by noise and overfitting problems. The random forest models are robust in handling missing values, noise. Such models are easy to use because of a small number of parameters required to be fitted to the given data set, as described in [4].

The imbalance in payment transaction data causes problems which could be partly resolved by using sampling methods. The main strategies are as follows: oversampling, undersampling, and synthetic generation of data. In the related literature it has been reported that the best performance is provided by the under-sampling strategy where the fraud rate is increased by designing a balanced class distribution, that is achieved by reducing a disproportion between the majority (normal) and minority (fraud) classes of payment transactions. In most cases, the imbalanced data make the specificity and sensitivity critical for fraud detection. An increase in sensitivity reduces the false negative outcomes while increasing the true positive detection [6].

Application of Hidden Markov Model transaction sequences are additional features in models where correlations improve the detection accuracy, as described in [14]. According to [27] a transaction aggregation and descriptive features about the past periods allow for a 28% increase in detection of fraudulent transactions.

Random Forest (RF) is a well-known ML technique which aggregates decision tree (DT) models built on a given data set. Within this technique DT models are randomised by using different combinations of explaining variables as well as by using different data samples as shown in [3].

This paper describes an experimental approach based on the RF and ML techniques to design solutions for detecting credit card frauds. Within this study the frauds are detected by analysing the client's purchasing behaviour. Finally the designed models are analysed and compared in terms of detection accuracy taking into account the fact that credit card data are imbalanced.

Related work

Machine Learning concepts have been efficiently used for detection of abnormal patterns [16, 15] and estimation of brain development [11, 25], trauma severity estimation and survival prediction [10, 21, 20], collision avoidance at Heathrow [22], brain computer interface [24] as well as in early detection of bone pathologies [1, 8].

The detection errors could be minimised by using a random walk sampler based on Markov chains, as described in [9, 10]. Such methods have provided reliable estimates of predictive posterior density distribution, which is critically important for evaluation and minimisation of risks in the presence of uncertainties [19, 23].

Methods and Data

The credit card data used in this study contains 284,315 legitimate and 492 illegitimate transactions, showing an imbalance rate 0.0017. It is therefore of crucial importance to find parameters of ML models which will provide the maximal detection accuracy on such an imbalanced data. In this study solutions developed within the RF framework have been explored with the following range of parameters: (i) the number of decision trees vary between 200 and 500, (ii) the number of explaining variables was set in the range 5 to 25.

Receiver Operator Characteristic (ROC) curve is a graphical plot used to show the binary classification outcomes. The ROC is constructed by plotting True positive against False Positive rates. The TP rate is a portion of data samples which are correctly detected as positive. The False Positive rate is a portion of incorrectly detected events to be positive being negative. The ROC curve depicts the trade-offs between sensitivity and specificity. A curve close to the top left corner of ROC indicates a better model performance. The closer the curve to the ROC diagonal line 50%, the less accurate the model. Evaluation of ROC is made within the area under the ROC curve and the higher the AUC the better for detection accuracy.

Results

The best performance in terms of confusion matrix has been achieved with a model providing the detection error 0.04 and true positive rate 0.194. Fig. 1 shows the fitting (training) error over the number of decision trees in the RF model.

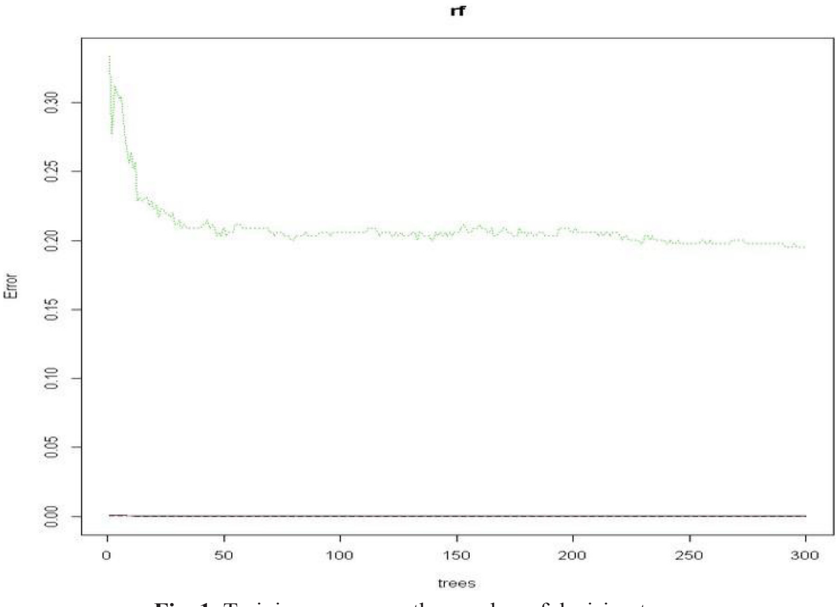

Fig. 1. Training error over the number of decision trees in the Random Forest model

We can see that the error rate drops down with the number of DT models and becomes stable after 200 trees. It is unlikely that the use of a larger number of DT models could further improve the detection accuracy.

Fig. 2 shows a distribution of the number of DT nodes in the designed RF model. The number largely varies from 80 to 160 nodes with a mean value around 130 nodes.

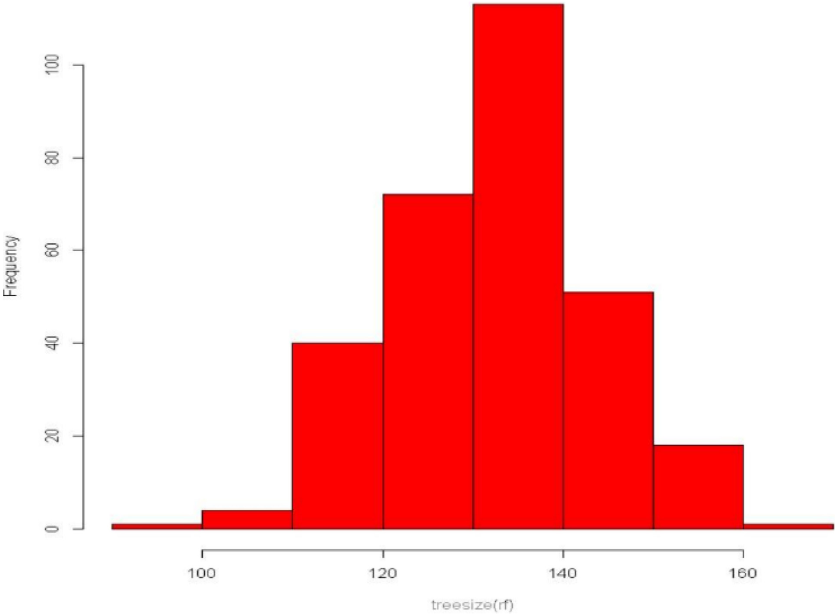

Fig. 2. Distribution of Decision Tree nodes in the designed Random Forest model

The performances of ML techniques applied to the imbalanced problem are critically dependent on parameters of models and imbalance strategies. The oversampling strategy increases the rate of positive transactions to the same value as in the negative (normal or legitimate) transactions. According to our observations shown in Table 1 the oversampling strategy has the smallest sensitivity, defined as True Positive rate, 0.7770. In contrast, the under-sampling strategy has the highest specificity 0.8851 whilst the sensitivity of the synthetic strategy is only 0.8446. Therefore, the under-sampling is more efficient when sensitivity is required to be maximal. The ability of ML techniques to detect the True Negative (legitimate) events, defined by specificity, can play the second role in fraud detection. The specificity is typically defined by a trade-off between the financial losses.

It can be seen in Table 1 that the oversampling strategy has a specificity of 0.9999 following after the synthetic strategy with a specificity 0.9890 and the under-sampling strategy with 0.9781.

The overall accuracy was highest in the oversampling strategy 0.9995, compared to 0.9781 of the under sampling and 0.9890 of the synthetic strategies. The accuracy however plays the secondary role in imbalance data problems.

Table 1.

Performances of the oversampling, undersampling, and synthetic strategies in terms of the detection accuracy, sensitivity, and specificity

Method	Accuracy	Sensitivity	Specificity
OverSample	0.9995	0.9779	0.9888
UnderSample	0.7770	0.8851	0.8446
Synthetic	0.9995	0.9781	0.9890

Discussion

The presented study aimed to experimentally test the performance of ML strategies known in the literature on a benchmark representing a real-world fraud detection imbalance problem. The use of a confusion matrix for evaluation of the detection performance cannot be informative because of the imbalance data. For these reasons the experimental results were analysed in terms of sensitivity and specificity.

The experiments have been conducted with Random Forest models which allow practitioners to use a small number of model's parameters. The experiments were run with the number of decision trees in the aggregated model as well as with another two parameters of aggregate randomisation, specifically data sample and feature sampling rates.

Conclusions

Across the world there is a growing need to prevent losses caused by frauds in payment transactions. Fraudulent transactions led to growing financial losses. Machine learning (ML) and Data Mining techniques have been shown to be efficient in improving the fraud detection accuracy in many applications described in the related literature.

This paper has proposed an experimental way of building ML solutions to the problem aimed at minimising the financial losses by monitoring the customer's behaviour of using credit cards. The design is tested on a public data set available for the research community in terms of detection accuracy.

In practice the ML fraud detection techniques require a retrospective transaction data showing the card holder's behaviour. There are a wide range of ML approaches used in credit card fraud detection. This study was mainly focused on Random Forest applied to a real credit card data set.

The RF models performed often better than the other ML techniques. When data are imbalanced special strategies are required to deal with the imbalance problem. Such applications as fraud detection require the analysis of True positive (sensitivity) and True negative (specificity) rates.

The study has shown that the use of the under-sampling strategy allows the RF models to achieve a greater sensitivity on the imbalance data. In practice the sensitivity plays the first role in planning efficient strategies of fraud detection in credit cards.

The use of the RF models for detecting credit card fraud transactions helps to analyse and provide practitioners with a new insight into data.

Acknowledgements

The authors are grateful to Dr Livija Jakaite and Dr Vitaly Schetinin from the University of Bedfordshire for insights in research and useful comments.

References / Список литературы

1. Akter, M., Jakaite, L. Extraction of texture features from x-ray images: Case of osteoarthritis detection. In: X.S. Yang, S. Sherratt, N. Dey, A. Joshi (eds.) Third International Congress on Information and Communication Technology, 2019, pp. 143–150. Springer Singapore, Singapore. https://doi.org/10.1007/ 978-981-13-1165-9 13
2. Bhattacharyya, S., Jha, S., Tharakunnel, K., Westland, J.C. Data mining for credit card fraud: A comparative study. *Decision Support Systems*, 2011, vol. 50(3), pp. 602– 613. https://doi.org/10.1016/j.dss.2010.08.008.
3. Breiman, L. Random forests. *Machine Learning*, 2001, vol. 45(1), pp. 5–32. https://doi.org/10.1023/A:1010933404324
4. Breiman, L., Friedman, J., Olshen, R., Stone, C. Classification and RegressionTrees. Chapman and Hall, 1984.

5. Correa Bahnsen, A., Aouada, D., Stojanovic, A., Ottersten, B. Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, 2016, vol. 51, pp. 134–142. https://doi.org/10.1016/j.eswa.2015.12.030
6. Dzakiyullah, N.R., Pramuntadi, A., Fauziyyah, A.K. Semi-supervised classification on credit card fraud detection using autoencoders. *Journal of Applied Data Sciences*, 2021, vol. 2(1), pp. 01–07.
7. Fiore, U., De Santis, A., Perla, F., Zanetti, P., Palmieri, F. Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, 2019, vol. 479, pp. 448–455.
8. Jakaite, L., Schetinin, V., Hladuvka, J., Minaev, S., Ambia, A., Krzanowski, W. Deep learning for early detection of pathological changes in x-ray bone microstructures: case of osteoarthritis. *Scientific Reports,* 2021, vol. 11. https://doi.org/10.1038/s41598-021-81786-4
9. Jakaite, L., Schetinin, V., Maple, C. Bayesian assessment of newborn brain maturity from two-channel sleep electroencephalograms. *Computational and Mathematical Methods in Medicine*, 2012, pp. 1–7. https://doi.org/10.1155/2012/ 629654
10. Jakaite, L., Schetinin, V., Maple, C., Schult, J. Bayesian decision trees for EEGassessment of newborn brain maturity. *The 10th Annual Workshop on Computational Intelligence UKCI 2010*. 2010. https://doi.org/10.1109/UKCI.2010.5625584
11. Jakaite, L., Schetinin, V., Schult, J. Feature extraction from electroencephalograms for Bayesian assessment of newborn brain maturity. *24th International Symposium on Computer-Based Medical Systems (CBMS),* 2011, pp. 1–6. https://doi.org/10.1109/CBMS.2011.5999109
12. Jha, S., Westland, J.C. A descriptive study of credit card fraud pattern. *GlobalBusiness Review,* 2013, vol. 14(3), pp. 373–384. https://doi.org/10.1177/0972150913494713
13. Kaggle: Credit card fraud detection. https://www.kaggle.com/mlg-ulb/creditcardfraud
14. Lucas, Y., Portier, P.E., Laporte, L., He-Guelton, L., Caelen, O., Granitzer, M., Calabretto, S. Towards automated feature engineering for credit card

fraud detection using multi-perspective hmms. *Future Generation Computer Systems*, 2020, vol. 102, pp. 393–402.

15. Nyah, N., Jakaite, L., Schetinin, V., Sant, P., Aggoun, A. Evolving polynomial neural networks for detecting abnormal patterns. *2016 IEEE 8th International Conference on Intelligent Systems (IS)*, 2016, pp. 74–80. https://doi.org/10. 1109/IS.2016.7737403
16. Nyah, N., Jakaite, L., Schetinin, V., Sant, P., Aggoun, A. Learning polynomial neural networks of a near-optimal connectivity for detecting abnormal patterns in biometric data. In: 2016 SAI Computing Conference (SAI), 2016, pp. 409–413. https://doi.org/10.1109/SAI.2016.7556014
17. Pourhabibi, T., Ong, K.L., Kam, B.H., Boo, Y.L. Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems*, 2020, vol. 133, 113303. https://doi.org/10.1016/j.dss.2020.113303
18. Prusti, D., Rath, S.K. Fraudulent transaction detection in credit card by applying ensemble machine learning techniques. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, 2019, pp. 1–6.
19. Schetinin, V., Jakaite, L. Extraction of features from sleep EEG for Bayesian Assessment of brain development. *PLOS ONE,* 2017, vol. 12(3), pp. 1–13. https://doi.org/10.1371/journal.pone.0174027
20. Schetinin, V., Jakaite, L., Krzanowski, W. Bayesian averaging over decision tree models: An application for estimating uncertainty in trauma severity scoring. *International Journal of Medical Informatics,* 2018, vol. 112, pp. 6-14. https://doi.org/10.1016/j.ijmedinf.2018.01.009
21. Schetinin, V., Jakaite, L., Krzanowski, W. Bayesian averaging over decision tree models for trauma severity scoring. *Artificial Intelligence in Medicine,* 2018, vol. 84, pp. 139–145. https://doi.org/10.1016/j.artmed.2017.12.003
22. Schetinin, V., Jakaite, L., Krzanowski, W. Bayesian learning of models for estimating uncertainty in alert systems: Application to air traffic conflict avoidance. *Integrated Computer-Aided Engineering,* 2018, vol. 26, pp. 1–17. https://doi.org/10.3233/ICA-180567
23. Schetinin, V., Jakaite, L., Krzanowski, W.J. Prediction of survival probabilities with Bayesian decision trees. *Expert Systems with Applica-*

tions, 2013, vol. 40(14), pp. 5466 – 5476. https://doi.org/10.1016/j.eswa.2013.04.009

24. Schetinin, V., Jakaite, L., Nyah, N., Novakovic, D., Krzanowski, W. Feature extraction with GMDH-type neural networks for EEG-based person identification. *International Journal of Neural Systems*, 2018. https://doi.org/10.1142/ S0129065717500642
25. Schetinin, V., Jakaite, L., Schult, J. Informativeness of sleep cycle features in bayesian assessment of newborn electroencephalographic maturation. *2011 24th International Symposium on Computer-Based Medical Systems (CBMS),* 2011, pp. 1–6. https://doi.org/10.1109/CBMS.2011.5999111
26. UK Finance: Fraud the facts 2019. https://www.ukfinance.org.uk/ policy-and-guidance/reports-publications/fraud-facts-2019
27. Whitrow, C., Hand, D.J., Juszczak, P., Weston, D., Adams, N.M. Transaction aggregation as a strategy for credit card fraud detection. *Data mining and knowledge discovery*, 2009, vol. 18(1), pp. 30–55.

DATA ABOUT THE AUTHORS

Courage Mabani, graduate student

School of Computer Science and Technology, University of Bedfordshire
Luton, LU1 3JU, United Kingdom
courage.mobani@study.beds.ac.uk
ORCID: https://orcid.org/0000-0002-1567-4302

Andrey A. Tuskov, candidate of economic sciences

Penza State University; K.G. Razumovsky Moscow State University of technologies and management (the First Cossack University)
40, Krasnaya Str., Penza, 440026, Russian Federation; 73, Zemlyanoy Val Street, Moscow, 109004, Russian Federation
tuskov@mail.ru
ORCID: https://orcid.org/0000-0003-1760-2676

Elizaveta V. Shchanina, student

Penza State University

40, Krasnaya Str., Penza, 440026, Russian Federation
shchanina03@list.ru

ДАННЫЕ ОБ АВТОРАХ

Мабани Кураж, аспирант

Школа компьютерных наук и технологий, Университет Бедфордшира
Лутон, LU1 3JU, Великобритания
courage.mobani@study.beds.ac.uk

Тусков Андрей Анатольевич, кандидат экономических наук

ФГБОУ ВО «Пензенский государственный университет»; Пензенский казачий институт технологий (филиал) МГУТУ им. К.Г. Разумовского
ул. Красная, 40, г. Пенза, 440026, Российская Федерация; ул. Земляной Вал, 73, г. Москва, 109004, Российская Федерация
tuskov@mail.ru

Щанина Елизавета Вячеславовна, студент

ФГБОУ ВО «Пензенский государственный университет»
ул. Красная, 40, г. Пенза, 440026, Российская Федерация
shchanina03@list.ru

Поступила 01.08.2022	Received 01.08.2022
После рецензирования 20.08.2022	Revised 20.08.2022
Принята 01.09.2022	Accepted 01.09.2022