As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 7(06): 1-3, 2021
Copyright © 2021International Journal for Modern Trends in Science and Technology
ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0706001
Available online at:http://www.ijmtst.com/vol7issue06.html

Fraud Detection Using Multi-layer Heterogeneous EnsembleMethod

Haritha Rajeev[1] | Neenu Kuriakose[2]

[1]Asst.Professor, Department of computer application, Kesari Arts And Science College, Kerala, India
[2]Research Scholar, Kerala, India

To Cite this Article
Haritha Rajeev,; Neenu Kuriakose. Fraud Detection Using Multi-layer Heterogeneous EnsembleMethod. *International Journal for Modern Trends in Science and Technology* 7, 1–3 (2021).

Article Info
Received on 27-April-2021, Revised on 20-May-2021, Accepted on 24-May-2021, Published on 29-May-2021.

ABSTRACT

Fraudulent detection is a large number of exercises that try to keep cash or property out of the way. Fraud surveillance is used in many businesses such as banking or security. At the bank, misrepresentation may involve producing checks or using a Credit Card taken. Different types of robberies can include misfortune or create a problem with the expectation of only a paid Layer Ensemble Method running other AI fields including collecting learning. Recently, there have been one deep group models deployed with a large number of classifiers in each layer. These models, as a result, require a much larger calculation. In addition, the deep integration models are available that use all the separating elements including the unnecessary ones that can reduce the accuracy of the group. In this experiment, we propose a multi-layered learning structure called the Two-Layer Ensemble System to address the issue of definition. The proposed framework is working with a number of weird filters to get the troupe jumper sity, in these lines being a technology in the use of equipment.

KEYWORDS: *Credit Card Fraud, MULES, gcForest*

I. Introduction

A credit card is a statement issued to a client to allow a cardholder to pay a merchant in transaction with a cardholder guarantee to the payer in order to pay it in spite of the following approved total amount. The sponsor of the card is usually a bank, The cardholder can get money from the merchant or as a forward share.

As credit card exchange rates increase, cash-related frauds each year as a result of fraudulent use of credit cards are increasing dramatically. It has changed into an important problem-based measurement problem. Visa impulse has cost about $ 21 billion worldwide and will look to move to $ 31 billion by 2020. There are various ways to deal with creating a Credit cardexchange. At the time when the card actually appears to make a Most ecommerce companies part, withdraw, or exchange. The following is the point where the card is lost, especially for purchases or divisions made to (requires certain tricks such as card verification value, cardholder name, PIN, security question). Credit card misconduct occurs when credit card data or a certain number of consent is taken and used without allowing them to defraud money, objects and organizations. Cybercriminals are basically not giving credit card fraud, they are constantly advancing sophisticated methods, so there is a real need to make advanced and dynamic systems designed to adapt to the clever flow of trading model distortions will be fraudulently monitored if the system detects customer deviations. A few data mining techniques are used to identify the problem of credit card fraud. This paper aims to design an effective classification-based study program that is inspired by representation reading using layer analysis by DNNs. To achieve this goal, the following objectives are explained:

• Create a compilation of a novel model that involves the storage of multiple layers with different partitions in each layer (called MULES). By combining the benefits of learning together with in-depth learning, the proposed system is expected to perform well in many classification tasks.

• Investigate how to simultaneously select subdivision elements and features in each layer. The proposed method therefore can overcome the limit of the existing algorithm for data mining

II. PROPOSED SYSTEM

Integration algorithm then trains in prediction. In fact, this model is based on the concept of DNNs where data is transferred to several components in a learning, and Development have researched on Two new, process. However, it is different from DNNs since the data in MULES is processed under a forwarding system where the information is transmitted only from one layer to the next layer and does not include back transmission as neural network channels. In addition, different learning algorithms will be used for each layer, while DNNs only work with layers of neurons. Image of MULES with 3 layer Architecture as shown in figure1. The results of the Decision Tree, Naïve Bayes, and Random Forest classifier in the first layer are combined with the actual training data to produce the second layer. The same scheme is applied in the second layer with two dividers (SVM and Decision Tree) and in the third layer with two dividers (LDA and SVM), before the release of the third layer combined with the final prediction.In integrated components, an algorithm is used to integrate the prediction of shared prediction dividers. In this study, we used the Sum Rule method of integration. The Sum Rule summarizes the predictions of each situation in relation to each class label and provides an example of a class label We show an example of the foundations of MULES found in the credit card database. For each layer, MULES has selected the appropriate dividers and their characteristics to produce input in the next layer. In detail, in the first layer

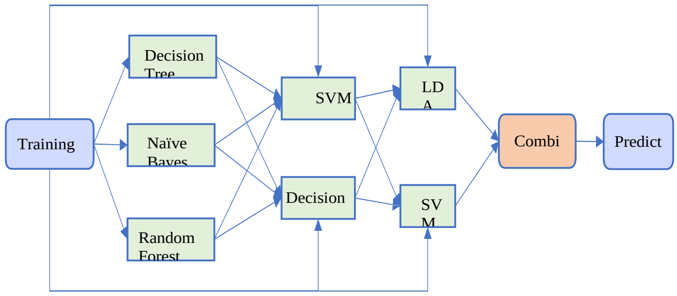


Fig 1:Flow chart of Proposed Model

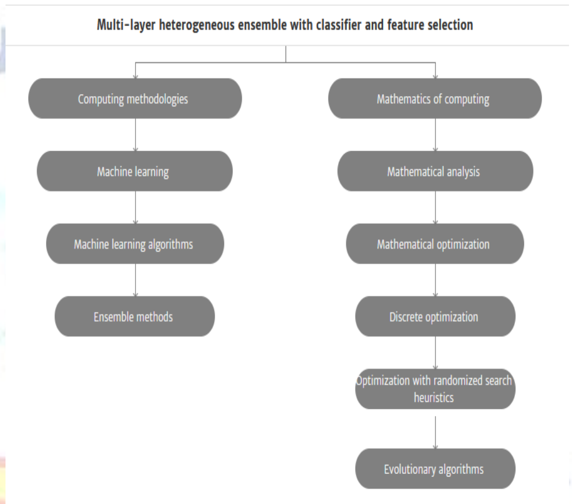
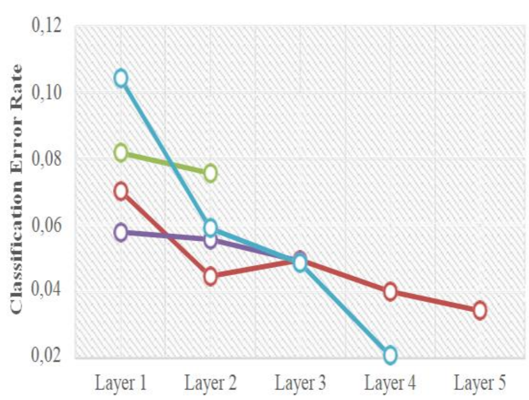


Fig2:Multi –layer heterogeneous ensemble with classifier and feature selection

In detail, in the first layer, MULES selected 3 divisions: Rands Rango, Decision Tree, and KNN. Each separator has used its selected features in the original features. In the second layer, classifiers outside the Decision Tree were selected with sets of various features found in the first and predictable features. In the third and fourth layers, two categories of KNN and Logistic Regression were selected with sets of various features. By selecting the appropriate subdivisions for each layer and the appropriate characteristics for each subdivision, MULES can achieve high predictive accuracy and efficiency in resource utilization depending on memory and computational requirement.



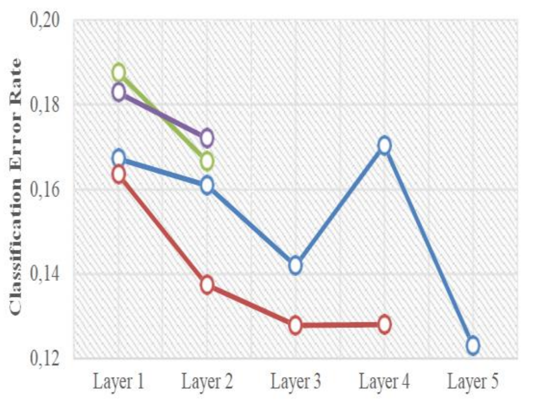


Fig3:The changes of classification error rate in each layer

III. Experimental Result

Classification time: Although MULES takes much higher train-ing time than gcForest, the classification time of MULES is lowerthan gcForest. On Credit card dataset, for example, MULES used3154.86 second for training process compared to only 311.78 ofgcForest. Meanwhile, gcForest used 0.62 second to classify all testinstances while MULES only used 0.26 second.

IV. Conclusion

In summary, we have introduced the Multi-Layer Heterogeneous Ensemble System (MULES) that is a layer-by-layer processing of DNNs. MULES incorporates several layers of a separate divider group where the sections train one train into new training details created by the previous layer. The new single-layer training data is the integration of the divider specs in the previous layer with the actual training data. . In the credit card database, for example, MULES spent 3154.86 seconds on the training process compared to only 311.78 gcForest. Meanwhile, gcForest used 0.62 seconds to separate all test conditions while MULES used only 0.26 seconds.

References

[1] MarijaBacauskiene, AntanasVerikas, AdasGelzinis, and Donatas Valincius. 2009. A feature selection technique for generation of classification committees and its application to categorization of laryngeal images. *Pattern Recognition* 42, 5 (2009), 645—65

[2] Boyuan Chen, Harvey Wu, Warren Mo, Ishanu Chattopadhyay, and Hod Lipson. 2018. Autostacker: A compositional evolutionary learning system. In *Proceedings of the Genetic and Evolutionary Computation Conference.* 402--409.

[3] Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acmsigkdd international conference on knowledge discovery and data mining.* 785—794

[4] Yijun Chen and Man Leung Wong. 2011. Optimizing stacking ensemble by an ant colony optimization approach. In *Proceedings of the 13th annual conference companion on Genetic and evolutionary computation.* 7—8

[5] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation* 6, 2 (2002), 182—197

[6] JanezDemšar. 2006. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research* 7, Jan (2006), 1—3

[7] Manuel Fernández-Delgado, Eva Cernadas, Senén Barro, and Dinani Amorim. 2014. Do we need hundreds of classifiers to solve real world classification problems? *The journal of machine learning research* 15, 1 (2014), 3133—3181

[8] Kyung-Joong Kim and Sung-Bae Cho. 2008. An evolutionary algorithm approach to optimal ensemble classifiers for DNA microarray data analysis. *IEEE Transactions on Evolutionary Computation* 12, 3 (2008), 377--388.

[9] Josef Kittler, Mohamad Hatef, Robert PW Duin, and Jiri Matas. 1998. On combining classifiers. *IEEE transactions on pattern analysis and machine intelligence* 20, 3 (1998), 226—239

[10] E. H. Budiarto, A. Erna Permanasari and S. Fauziati, "Unsupervised Anomaly Detection Using K-Means, Local Outlier Factor and One Class SVM," *2019 5th International Conference on Science and Technology (ICST)*, Yogyakarta, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICST47872.2019.9166366.

[11] H. Chen, S. Zhao, H. Bao and H. Kang, "Research of local outlier mining algorithm based on spark," *2017 First International Conference on Electronics Instrumentation & Information Systems (EIIS)*, Harbin, 2017, pp. 1-4, doi: 10.1109/EIIS.2017.8298559.

[12] B. Yu, M. Song and L. Wang, "Local Isolation Coefficient-Based Outlier Mining Algorithm," *2009 International Conference on Information Technology and Computer Science*, Kiev, 2009, pp. 448-451, doi: 10.1109/ITCS.2009.230.

[13] BaozhiQiu, Chenke Jia and Junyi Shen, "Local Outlier Coefficient-Based Clustering Algorithm," *2006 6th World Congress on Intelligent Control and Automation*, Dalian, 2006, pp. 5859-5862, doi: 10.1109/WCICA.2006.1714201

[14] M. Sakr, W. Atwa and A. Keshk, "Sub-Grid Partitioning Algorithm for Distributed Outlier Detection on Big Data," *2018 13th International Conference on Computer Engineering and Systems (ICCES)*, Cairo, Egypt, 2018, pp. 252-257, doi: 10.1109/ICCES.2018.8639409.

[15] W. Huang, D. Wu and J. Ren, "An Outlier Mining Algorithm in High-Dimension Based on Single-Parameter-k Local Density," *2009 Fourth International Conference on Innovative Computing, Information and Control (ICICIC)*, Kaohsiung, 2009, pp. 1192-1195, doi: 10.1109/ICICIC.2009.98.