AIRO JOURNALS

Free / Unpaid
Peer Reviewed
Multidisciplinary
International

ISSN:2320-3714
Volume:3 Issue:3
September 2022
Impact Factor:5.7
Subject: Engineering

A STUDY OF DEEP LEARNING APPROACHES FOR DETECTING FINANCIAL FRAUD

Mr. Pramod Salunkhe[*1], Dr. Ashish Chaurasia[*2]

[*1]Research Scholar, Department of Computer Science University of Technology, Jaipur, India
[*2] Professor, Department of Computer Science University of Technology, Jaipur, India

DECLARATION: I AS AN AUTHOR OF THIS PAPER /ARTICLE, HERE BY DECLARE THAT THE PAPERSUBMITTED BY ME FOR PUBLICATION IN THIS JOURNAL IS COMPLETEY MY OWN PREPARED PAPER. I HAVE CHECKED MY PAPERTHROUGHMY GUIDE/SUPERVISOR/EXPERT AND IF ANY ISSUE REGARDING COPYRIGHT/PATENT/ PLAGIARISM/ OTHER REAL AUTHOR ARISE, THE PUBLISHER WILL NOT BE LEGALLY RESPONSIBLE. IF ANY OF SUCH MATTERS OCCUR PUBLISHER MAY REMOVE MYCONTENT FROM THE JOURNAL

ABSTRACT

Frauds are said to be dynamic and devoid of patterns, making them challenging to detect. Fraudsters profit from recent technological improvements. Security precautions were overcome, causing a loss of millions of dollars. Using data mining techniques to monitor and spot unusual behavior is one way to track fraudulent transactions. Transactions. This article compares deep learning methods including auto encoders, convolutional neural networks, restricted Boltzmann machines, and deep belief networks to k-nearest neighbor (KNN), random forest, and support vector machines (SVM) (DBN). This study will make use of the European (EU), Australian, and German databases. The Area Under the ROC Curve (AUC), Matthews Correlation Coefficient (MCC), and Cost of Failure are the three assessment metrics that would be used.

Keywords: Deep learning, Fraud Detection, Machine learning;

INTRODUCTION

INTRUSION detection (ID) is a security solution for computers and networks. An ID System collects and analyses data from various areas of a device or network in order to spot potential security risks like intrusions (outside-the-board assaults) and harassment. ID employs a method called vulnerability assessment (scanning) to evaluate the safety of computer networks or systems. It benefits from an analysis of vulnerabilities. IDS, or intrusion detection system, stands for. For instance, the lock system in the house guards against theft. Additionally, the burglar alarm alerts the owner if someone tries to break into the house or breaks the lock. In order to stop the firewall from activating, firewalls also filter incoming Internet traffic very effectively. Firewalls won't notice external users connecting to the intranet via a modem on the company's private network, for instance. The misuse of a profitable company's system without triggering legal repercussions is referred to as "fraud." In a market where competition is fierce, fraud could become a business-critical issue if it is pervasive and preventive measures are unsafe. As a component of overall fraud protection, fraud detection automates or helps to lessen the manual portions of the scanning but testing process. In the data mining business and government, it is a well-known application. It's impossible to know for sure whether or not a request or transaction is legitimate. In reality, using mathematical algorithms to remove possible evidence of data fraud is the most cost-effective option.

REVIEW OF LITERATURE

G. S. Temponeras et al. [2019]the identification and dependability of FFS is one of the most significant financial problems. To identify the issues directly linked to FFS, many model machine training models were created. A novel model for detecting fraud using a deep, ANN is given. An experimental test of a new prediction model was performed utilizing data from Greek businesses in particular. The findings show that the suggested system is both robust and promising.

Salvatore Carta et al. [2019] proposes a new data intelligence tool built on the Multi Consensus Prudential model that evaluates the effectiveness of various cutting-edge categorization algorithms using both probabilistic and majority-based metrics. Regardless of data discrepancies, the objective is to increase the model's accuracy in spotting fraudulent transactions. A sizable real-world dataset was used to validate our model, which shows that it outperforms current state-of-the-art solutions in terms of ensemble models and classification techniques.

Xinwei Zhang et al. [2019] Existence of a sophisticated feature-engineered homogeneity-driven behavioral analysis-based fraud detection system (HOBA). We conduct a systematic analysis based on the actual experience of one of China's major commercial banks to gauge the effectiveness of the proposed approach. The test results show that our suggested strategy is a reliable yet feasible means of identifying credit card fraud. Practically speaking, because our methodology has a low false positive rate compared to other benchmarking methods, it may be able to identify relatively fraudulent transactions. Credit card issuers may utilize technique to identify fraudulent transactions efficiently, safeguard consumer interests, limit fraud, and reduce regulatory costs as a result of our study.

A. Kim et al. [2019] Consider how deep learning might influence financial risk management decisions. To predict how independent distribution dealers would protect future profits, we developed a deep learning algorithm. Traditional risk and behavior difficulties are part of the job. Traditional machine learning relies on the creation, upkeep, and analysis of specific functionalities, which may be costly. It also uses data that is indicative of the functionality-objective connection. As a result, it can be difficult to model complex, broad, and variable models like trader behavior. The answer is deep thought. Despite being avoided, mechanical functional engineering was more adaptable to changes due to its generative characteristics that clearly defined the goal. The outcomes of a comprehensive risk operating network demonstrate deep learning's application potential, provide guidance for designing network architectures, and demonstrate deep learning's superiority over machine learning or expectations based on rules.

EXPERIMENTATION

A. Data Sets

Our experiments will be performed on three data sets in this research. In September 2013, the European Dataset includes the transactions performed by credit card users over two days. Except for time and amount, all of the fields have been PCA converted. There are just 492 fraud cases out of a total of 284,807 fraud cases.

The Australian and German Datasets were both obtained from the UCI ML repository. There is no personal information in the data sets since they have been anonymized.

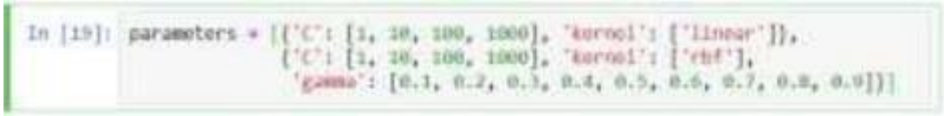
In Australia, there are 307 fraud cases and 383 regular incidents. In parallel, there are 1000 cases in the German data set, of which 300 are fraud cases and 700 are regular ones. The European dataset is significantly larger than the sum of the Australian and German datasets.

We make an effort to evaluate the effectiveness of various machine learning and deep learning models using data sets that range in size and complexity.

EXPERIMENTAL SETUP

- The work was all finished in Python, with libraries like NumPy, Pandas, Keras, Scikit-Learn, and Tensorflow being utilized. Rstudio was utilized to clean information now and again.
- We use cross-approval on the preparation informational index to find the ideal worth of the neighbors K for every informational index in K-closest neighbor. From that point forward, the best K for every informational index is used to play out extra investigation all in all assortment.

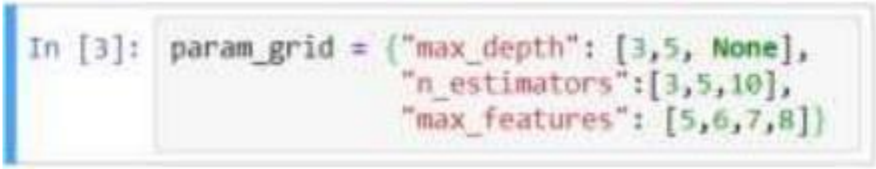
To recognize the ideal boundary for each model, we use a network based scan technique for Help vector machines and Irregular Backwoods. We use the Python technique GridSearchCV, involving the settings for the SVM and Irregular backwoods displayed in Figures 1 and 2.



```
In [15]: parameters = [{'C': [1, 10, 100, 1000], 'kernel': ['linear']},
                       {'C': [1, 10, 100, 1000], 'kernel': ['rbf'],
                        'gamma': [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]}]
```

Fig. 1. Parameters for SVM

To assess the models using all of the data, the best parameters are utilized.



```
In [3]: param_grid = {"max_depth": [3,5, None],
                      "n_estimators":[3,5,10],
                      "max_features": [5,6,7,8]}
```

Fig. 2. Parameters for Random Forest

The Autoencoders' fundamental concept is that they can recreate their input. As a result, we exclusively train the autoencoders on regular transactions for fraud detection. The experiment would generate reconstruction mistakes for each of the instances when running on the test data. Typical exchanges are supposed to have less remaking mistakes, while fraud exchanges/examples

ISSN:2320-3714
Volume:3 Issue:3
September 2022
Impact Factor:5.7
Subject: Engineering

are supposed to have more prominent qualities. A particular edge esteem is determined, and in the event that the remaking mistake surpasses it, the occurrence/exchange is viewed as fraudulent. In any case, the exchange is considered. We analyze different limit levels in our examinations and give the discoveries.

The confined Boltzmann machine (RBM), like Autoencoders, produces free energy that is then contrasted with an edge to distinguish ordinary versus fraudulent exchanges. Weiman Wang fostered the RBM model used here for fraud detection.

We use a changed rendition of AlbertUP's model for deep conviction organizations, which is executed in Tensorflow for regulated and solo example acknowledgment errands.

Rather than a 1D exhibit, we use CNN to change the dataset into a 2D cluster. The information is gone through a progression of convolutional and max-pooling layers prior to being leveled by a layer. At the SoftMax layer, the information is at last sorted. The architecture of the CNN we employ is shown in Figure 3.

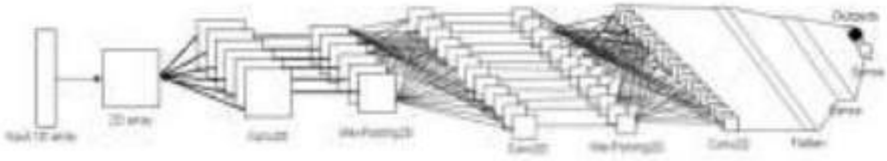


Fig. 3. CNN Architecture

We use cross-validation to evaluate our models and then use majoring voting to merge the best three performers. The model's fundamental structure is shown in the figure below.

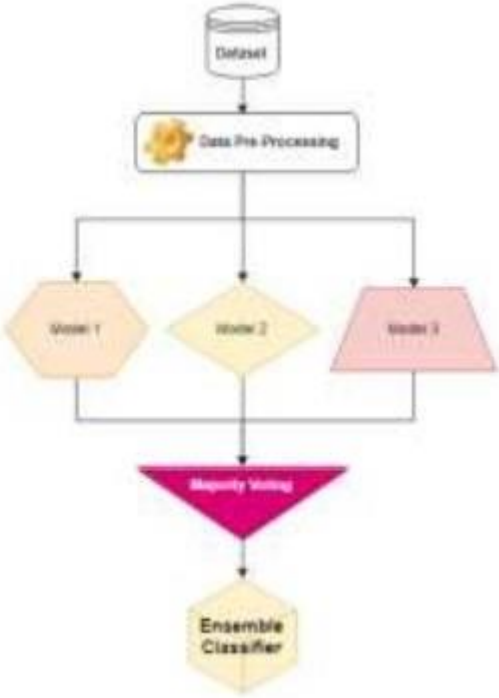


Fig. 4. Majority Voting Based Model Structure

C. Metrics for Evaluating The major assessment factors we examine in this research are listed below.

The Matthews Connection Coefficient is a measurement for deciding the nature of a double or two-class classifier. In 1975, Cerebrum W. Matthews proposed it. An immaculate gauge yields a worth of +1, though an irregular conjecture yields a worth of nothing. The phi coefficient is one more name for Matthews Connection. MCC is a far superior measurement than exactness or F1 score, as indicated by Davide Chicco, since the other two might be underhanded on the grounds that they don't consider each of the four components of the disarray grid [14].

The collector working trademark is addressed by the ROC bend. On account of the dataset's lopsidedness, it supports surveying the model's precision. The TPR on the x pivot is displayed against the FPR on the y hub in a ROC bend. At the point when two ROC bends have an equivalent Region under the bend (AUC), we really want to go further into the better boundaries, such the Expense of Disappointment. The concept behind the cost of failure is that each False Negative (Frauds identified as normal) costs $1000 to the company/entity, while False Positives (Normal occurrences reported as fraud) costs $100. We utilize this approach to assess the top three models since their MCC and AUC values are often quite close. Similarly, the cost of the ensemble classifiers that result is computed.

RESULTS

The results of the tests on the European Dataset, the Australian Dataset, and the German Dataset are shown below.

A. European Dataset

TABLE I. EUROPEAN DATASET RESULTS

Method			
	MCC	*AUC*	*CostofFailure*
RBM	0.176	0.9109	227360
Autoencoders	0.2315	0.8943	127220
RandomForest	0.7947	0.8507	30340
CNN	0.8096	0.8764	25700
SVM	0.8145	0.9004	21220
KNN0.83540.8887			22660
Ensemble(KNN, SVMandCNN)	0.82260.8964	21740	

The discoveries from the European Dataset are summed up in Table I. For an assortment of machine learning models, the table gives the Matthew Connection Coefficient (MCC) and the Region under the Bend measure (AUC).

Since RBM and AE have a high misleading positive rate (deception rate), they perform seriously with regards to MCC and cost. Irregular Woods has a high AUC and MCC. With regards to MCC and AUC, CNN, SVM, and KNN have the best outcomes. The SVM has the least expense of disappointment, while the Autoencoders and RBM have the most. Arbitrary woodland, despite the fact that creating great results, is wasteful with regards to cost. SVM, KNN, and CNN are the best three models for this informational index.

The best three performing models are combined to create the majority voting classifier. The ensemble approach outperforms SVM and CNN on their own, although it costs about the same as SVM. SVM, on the other hand, has a higher AUC value. If the business wants to save costs as much as possible, the suggestion is to use SVM instead of ensemble since the ensemble technique takes longer in terms of both training and testing, while SVM takes the least amount of time in terms of testing and training.

B. Australian Dataset

TABLE II. AUSTRALIAN DATASET RESULTS

Method	*MCC*	*AUC*	*CostofFailure*
RBM	0.15	0.5546	24600
Autoencoders	0.2318	0.6174	12220
CNN	0.6408	0.8227	6430
RandomForest	0.684	0.8416	4700
KNN	0.6905	0.8425	6460
DBN0.69990.8441			6790
SVM0.7085		0.8551	3380
Ensemble1(KNN, SVM,DBN)	0.7144	0.8573	5290
Ensemble2(KNN,SVM, Random Forest)	0.7281	0.8655	3470

The findings from the Australian Dataset are summarized in Table II. RBM and AE have the least exhibition of the multitude of methods, as can be shown. Regarding AUC and MCC, SVM, DBN, and KNN are awesome. In contrast with the others, Arbitrary Woodland and CNN have magnificent qualities. Two group models were picked. The KNN, DBN, and SVM classifiers make up the primary outfit model (Troupe 1). The models with the most minimal expense of disappointments are utilized to make a subsequent outfit (Troupe 2): KNN, SVM, and Irregular

Backwoods. The greatest expense of disappointment is RBM and AE, which have countless bogus up-sides inferable from the edge.

Table II demonstrates that contrasted with single SVM and different procedures, Group 1 (KNN, SVM, DBN) expanded MCC and AUC execution. Nonetheless, it is more costly than irregular timberland and SVM. This is on the grounds that to the significant expense of disappointment upsides of KNN and DBN, which no doubt affected the classification. Outfit 2 (KNN, SVM, Irregular woods), then again, got a superior MCC, AUC, and cheaper worth by incorporating classifiers with the smallest expense of disappointment methods. We can see that joining methods has brought about superior generally speaking results, and Group 2 is the best method of all since it has the best MCC, AUC, and cost.

C. German Dataset

TABLE III. GERMAN DATASET RESULTS

Method	*MCC*	*AUC*	*CostofFailure*
RBM	0.0984	0.5524	14160
Autoencoders	0.139	0.5614	22640
KNN	0.2487	0.6047	21100
DBN	0.2725	0.5873	23640
RandomForest	0.2912	0.6437	16970
SVM0.40380.6857			16400
CNN0.4291		0.7056	14220
Ensemble(SVM, CNN, RandomForest)	0.4439	0.7011	15620

The findings for the German Dataset are summarized in Table III. When we look at the table's findings, we can see that the top models in terms of performance are SVM, Random Forest, and CNN (AUC and MCC values). Random forest, CNN, and SVM have lower failure costs than other models. As a result, the majority voting classifier is built using an ensemble of these three models.

In general, the findings shown in tables I, II, and III for the three data sets studied indicate that combining the best models outperforms using single models. On smaller data sets, the ensemble improvement is more noticeable (the German dataset and the Australian dataset). It provides results that are somewhat less than SVM for the European dataset.

For smaller datasets, Random Forest is the best option. Convolutional Neural Networks were discovered to be the best deep learning technique since they provide excellent results for both the European and German datasets, but their performance on the Australian dataset was fourth best and their cost of failure was comparable to KNN. It also had the lowest cost for the German dataset. For all datasets, Table IV highlights the frequency of individual models placing in the top three performing models. SVM was consistently one of the top models across all data sets. KNN also provides excellent results with both big and small datasets.

TABLE IV. TOP PERFORMING MODELS

Method	*NumberoftimesinTop3*
SupportVectorMachines3Times	
K-NearestNeighbors2Times	
ConvolutionalNeuralNetworks2	Times
RandomForest2Times	
DeepBeliefNetwork1Time	

CONCLUSION

Fraud detection research has been happening for very nearly 20 years, and it has used various procedures, from manual check to client end verification. In this space, machine learning models have likewise had a great deal of progress. Deep learning models have recently been utilized in various applications, attributable to expanded PC ability and lower figuring costs.

This article presents an experimental assessment of deep learning models for the recognizable proof of fraudulent exchanges utilizing different informational collections. The essential objective of this exploration is to figure out which procedures are the most ideal for various sorts of datasets. Since numerous organizations are putting resources into imaginative ways of improving their main concerns nowadays, this study might help professionals and organizations better comprehend how different techniques perform on various datasets.

SVMs, maybe combined with CNNs for a more steady exhibition, are the best procedures for distinguishing fraud with greater datasets, as per our exploration. SVM, Irregular Woodland, and KNN group techniques might offer phenomenal upgrades for more modest datasets. Convolutional Brain Organizations (CNN) outflanks other deep learning methods like Autoencoders, RBM, and DBN generally speaking.

REFERENCES

1. G. S. Temponeras, S. N. Alexandropoulos, S. B. Kotsiantis, and M. N. Vrahatis, "Financial Fraudulent Statements Detection through a Deep Dense Artificial Neural Network," 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA), PATRAS, Greece, 2019, pp. 1-5. doi: 10.1109/IISA.2019.8900741.
2. Salvatore Carta, Gianni Fenu, Diego ReforgiatoRecupero, Roberto Saia, ―Fraud detection for Ecommerce transactions by employing a prudential Multiple Consensus model‖, Journal of Information Security and Applications 46 (2019) 13–22.
3. EunjiKima, JehyukLeea, HunsikShina, HoseongYanga, SungzoonChoa, Seung-Kwan Namb, Youngmi Song b, Jeong-a Yoonb, Jong-il Kim, ―Champion-challenger analysis for credit card fraud detection: Hybrid ensemble and deep learning‖, Expert Systems With Applications 128 (2019) 214–224.
4. Zhang, X., Han, Y., Xu, W., & Wang, Q. (2019). HOBA: A Novel Feature Engineering Methodology for Credit Card Fraud Detection with a Deep Learning Architecture. Information Sciences. doi:10.1016/j.ins.2019.05.023.
5. Kim, A., Yang, Y., Lessmann, S., Ma, T., Sung, M.-C., & Johnson, J. E. V. (2019). Can Deep Learning Predict Risky Retail Investors? A Case Study in Financial Risk Behavior Forecasting. European Journal of Operational Research. doi:10.1016/j.ejor.2019.11.007.

6. A. M. Mubalaike and E. Adali, "Deep Learning Approach for Intelligent Financial Fraud Detection System," 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, 2018, pp. 598-603. doi: 10.1109/UBMK.2018.8566574.

7. S. Panigrahi, A. Kundu, S. Sural, and A. K. Majumdar, "Credit card fraud detection: a fusion approach using dempstershafer theory and bayesian learning," *Information Fusion*, vol. 10, no. 4, pp. 354–363, 2009.

8. V. Sharma et al., "Cooperative trust relaying and privacy preservation via edge-crowdsourcing in social Internet of Things," *Generation Computer Systems*, 2017.

9. S. Vishal et al., "Computational offloading for efficient trust management in pervasive online social networks using osmotic computing," *IEEE Access*, vol. 5, pp. 5084–5103, 2017.

10. S. Vishal, Y. Ilsun, and K. Ravinder, "Isma: Intelligent sensing model for anomalies detection in cross platform osns with a case study on iot," *IEEE Access*, vol. 5, pp. 3284–3301, 2017.V

11. V. Wyk and Hartman, "Automatic network topology detection and fraud detection," U.S. Patent No. 9,924,242. 20 Mar. 2018.

12. A. Favila and P. Shivam, "Systems and methods for online fraud detection," U.S. Patent Application No. 15/236,077, 2018.

13. K. RamaKalyani and D. UmaDevi, "Fraud detection of credit card payment system by genetic algorithm," *International Journal of Scientific & Engineering Research*, vol. 3, no. 7, 2012.

14. Chang, W.-H., & Chang, J.-S. (2012). An effective early fraud detection method for online auctions. Electronic Commerce Research and Applications, 11(4), 346–360. doi:10.1016/j.elerap.2012.02.005.

15. A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams and P. Beling, "Deep learning detecting fraud in credit card transactions," 2018 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, 2018, pp. 129-134. doi: 10.1109/SIEDS.2018.8374722.