Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

Deep Convolution Neural Network Model for Credit-Card Fraud Detection and Alert

Joy Iong-Zong Chen,
Professor,
Department of Communication Engineering,
Da-Yeh University,
Chang-Hua, Taiwan
Email id: jchen@mail.dyu.edu.tw

Kong-Long Lai,
Department of Electrical Engineering,
Da-Yeh University,
Chang-Hua, Taiwan

Abstract: With the exponential increase in the usage of the internet, numerous organisations, including the financial industry, have operationalized online services. The massive financial losses occur as a result of the global growth in financial fraud. Henceforth, devising advanced financial fraud detection systems can actively detect the risks such as illegal transactions and irregular attacks. Over the recent years, these issues are tackled to a larger extent by means of data mining and machine learning techniques. However, in terms of unknown attack pattern identification, big data analytics and speed computation, several improvements must be performed in these techniques. The Deep Convolution Neural Network (DCNN) scheme based financial fraud detection scheme using deep learning algorithm is proposed in this paper. When large volume of data is involved, the detection accuracy can be enhanced by using this technique. The existing machine learning models, auto-encoder model and other deep learning models are compared with the proposed model to evaluate the performance by using a real-time credit card fraud dataset. Over a time duration of 45 seconds, a detection accuracy of 99% has been obtained by using the proposed model as observed in the experimental results.

Keywords: Artificial neural network, deep learning, financial fraud, convolution neural network, data mining

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

1.Introduction

Over the recent past, the number of businesses, online services, and internet users has exploded. In recent years, everyone has utilized internet banking systems for cash transfers, debit and credit cards for shopping, and online bill payment services [1]. This technology makes our lives easier with various benefits like shopping cash-free, avoiding long queues while paying bills or purchasing tickets and so on [2, 3]. However, financial frauds and unauthorized payments are causing large risks despite the positive aspects involved in online transactions. Money laundering, payment card frauds, insurance frauds, online identity theft, banking transaction fraud and many financial frauds are faced by the internet and online banking users [4]. Identification of the fraudulent financial activities are challenging due to its complications and sophistication. With the developments in modern technology, the financial frauds are also increasing significantly. Fraudulent accounts, scamming, phishing, documents falsification, fraudulent loans, credit card fraud, and online banking fraud are some of the several frauds faced by the financial systems [5]. On a yearly basis, several million dollars are lost due to fraud crimes that occur in financial establishments. The confidence of the customers and the financial situation of the establishment are highly affected [6].

Multiple financial frauds cause massive losses to companies and global financial institutions [7]. Without the knowledge of the authenticated user, unauthorized credit and debit card transactions, credit card theft and several other such fraudulent activities are alarming the world governments, clients and banking sector [8]. The financial fraud detection systems have the ability to identify unusual attacks and unauthorized access. These fraud detection mechanisms are constantly updated by financial institutions. These issues are addressed by data mining and machine learning tools that are commonly used over the last few years [9]. Various research literatures have proposed optimal solutions by using these tools and techniques. However, the integration of big data, memory cost and computational cost may still be improved by using these techniques to meet the requirements of the growing financial sector [10]. The major challenges addressed in financial fraud detection schemes are the constantly varying fraudulent behavior, lack of fraud transaction information tracking

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

mechanism, limitations of machine learning algorithms and other existing models and algorithms that are hard to train with the highly skewed financial fraud datasets.

2. Related Works

Significant legal costs, large investment losses and overthrow of entire organization has been caused due to huge frauds in businesses and global financial services [11]. In order to detect these frauds in financial sector, academic researches and investors have invested considerable efforts and proposed various technological solutions. The financial activities are largely benefited by fraud detection schemes. During the last decade, the fraud detection schemes have gained substantial research attention. In order to predict and detect the fraudulent financial activities, several deep learning and machine learning algorithms are exploited by various researchers. Credit card fraud detection using neural networks is one such example [12, 13]. A real dataset is used for credit card fraud detection and the performance of logistic regression and artificial neural network (ANN) has been compared in another study [14]. With the use of the training data, both models have exhibited similar performance according to the empirical result of the study. However, when test data is considered, ANN has performed better when compared to logic regression [15]. In order to expect an optimal output, it is essential to train the ANN with a real time dataset. Detection of tasks with abnormal behavior or fraud detection may not be performed by task classification of ANN if the model is not trained.

In order to detect credit card fraud, appropriate attributes are derived while enhancing the performance of the transaction capturing strategy using logistic regression [16]. The existing credit card fraud detection schemes may be improved with real time credit card transaction dataset. Using this dataset, the credit card frauds are identified based on location, transaction type product type and other significant classification parameters. However, the outcomes may be expected in a single category for fraud detection using logistic regression algorithm, which is a major drawback of this technique. This algorithm is also prone to over-fitting issues [17]. Fuzzy Darwinian system, fuzzy neural network, hidden Markov model, artificial immune system, k nearest neighbor, genetic algorithm, support vector machine, Bayesian network,

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

neural network and decision tree based credit-card fraud detection models have been compared.

The performance of these models are evaluated in terms of cost, speed and accuracy metrics and compared. In terms of processing speed, the hidden Markov model offered highest performance while in terms of accuracy, the fuzzy Darwinian method was optimal [18]. The classification issue may be addressed with great accuracy and variable cost of misclassification using an efficient algorithm based on a hybrid approach. High cost and low detection speed are some of the drawbacks of the fuzzy Darwinian approach despite its high performance [19, 20]. Low accuracy, high cost and large data are not scalable despite the quick processing speed of the hidden Markov model.

3. Proposed Work

The proposed financial fraud detection model is categorized into three stages based on before, during and after applying the model. Figure 1 represents the proposed model and its corresponding stages.

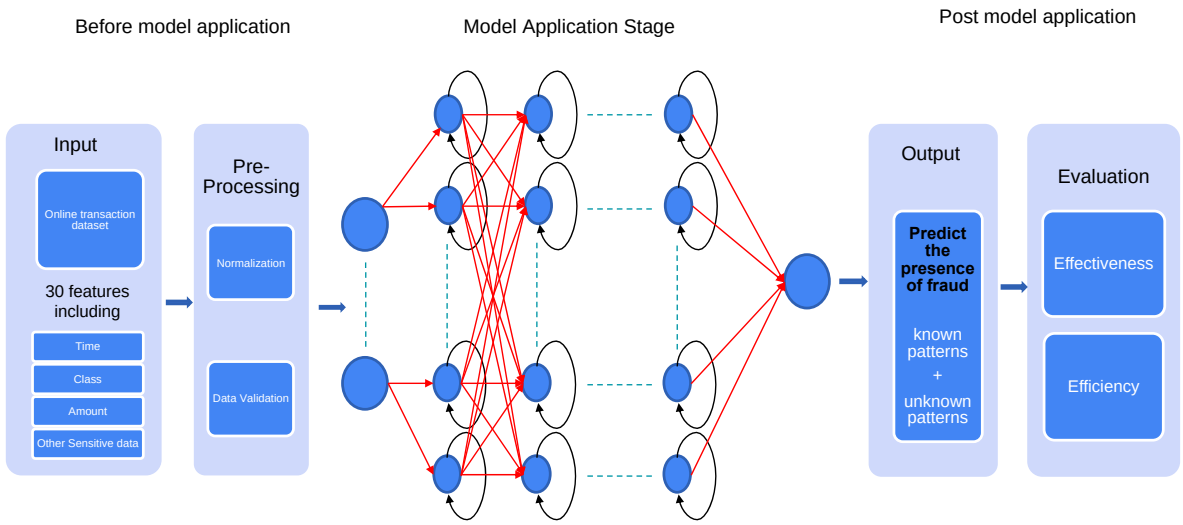

Fig. 1. Proposed Model

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

In order to train and test the model, the dataset and samples are prepared at the pre-application stage. A random sample of 5 million transactions over a duration of 24 hours is considered in which 6223 frauds exist. The dataset is strongly unbalanced and the fraudulent activity is at 0.12% of the overall transactions. Class, amount, time and other 30 subtle features are considered in the dataset. Various researchers have benchmarked these features for several studies. The dataset has to be validated, normalized and divided before the application of the model. Negative amount, empty values, negative time parameters in the dataset are checked in order to validate the data. The variables are rescaled in a range of -1 to 1 in order to enhance the accuracy of the result. This process is termed as data normalization. Further, the samples are divided into 70% and 30% for the purpose of training and testing respectively.

Once the pre-processing is performed, the memory based deep learning neural network, Convolution Neural Network technique is applied for the purpose of financial fraud detection. The historical input data and prediction results are compared based on the prior knowledge that enhances the prediction level. The sequential prediction issue is enabled through long dependency based learning using the DCNN architecture. Long memories and long-time patterns may be maintained using this model. Data can be retained over a long duration of time as a default behavior of the model. Thus detection and prediction is done efficiently using this model. The pre-processed data is fed to the model using the DCNN memory cell layers. The results are improved by the hyperparameters used by this model. The matrix, epochs, batch size and optimizers of the proposed model are estimated to evaluate its performance. On processing, the output and its evaluation is obtained during the final post model application stage. The algorithm makes the decision if the transaction is fraudulent or not and various parameters are tested to validate the authenticity of the output.

4. Results and Discussion

The time between multiple transaction, amount involved in the transaction and other sensitive features are analyzed in the data. An accuracy of 99% is received in 583 seconds with a loss rate of 0.32%. RMSprop, Adagrad and Adam optimizers are used for comparison and evaluation of various iterations and layers. With 3 encoding and decoding layers and an

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

iteration of 500 and 1000, we have obtained the training and validation results for accuracy and loss as shown in figure 2, 3, 4 and 5. The results obtained in terms of speed and accuracy from each optimizer is different due to its properties, varied iterations and layers. Best performing optimizer is identified from the experiments. From this comparison, we have observed that, an optimum accuracy is obtained with the help of Adam optimizer. It is also computationally efficient, requires lesser memory space and can handle big data.

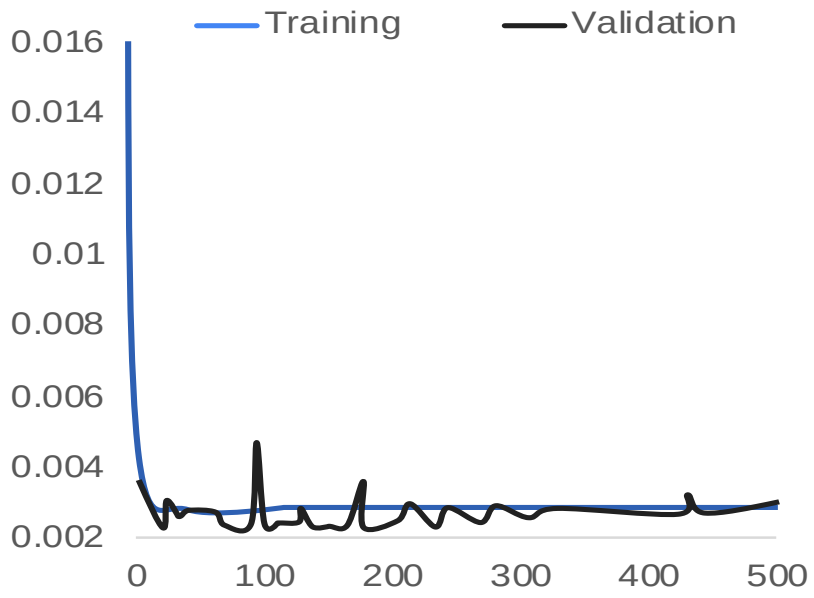

Fig. 2. Loss for 500 iterations

RNN cells are applied for efficient and effective prediction of financial frauds and credit-card scams. The result may be improved significantly with the appropriate choice from a wide variety of optimizers. From the experimental results, it is evident that, the number of iterations and layers does not show any prominent variation in the output. The time delay and loss rates are however affected by the variation in the number of layers and iterations. The correlation between the iterations and accuracy is significant as observed from the results. The number of iterations and loss rates are inversely proportional to each other. The loss rates decreases with the increase in accuracy leading to a negative correlation between the loss rates and accuracy.

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

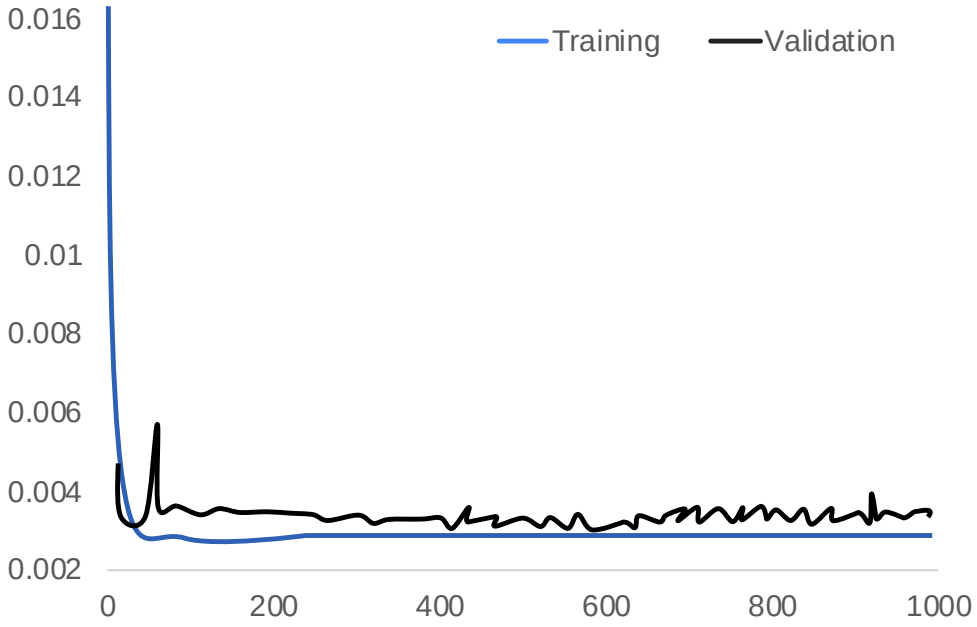

Fig. 3. Loss for 1000 iterations

Support Vector Machine (SVM), Logistic Regression (LR), and Random Forest (RF) based machine learning models are compared with the proposed DCNN model for the purpose of detection of financial fraud. The same data set is used for evaluation of these techniques. From the comparison results, it is evident that the execution, training and prediction time taken by RF is higher than that of DCNN despite similarity in accuracy. This reduces the effectiveness of the algorithm in fault detection. Accuracy and speed of SVM is comparatively lesser than DCNN making it unsuitable for complex and voluminous data.

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

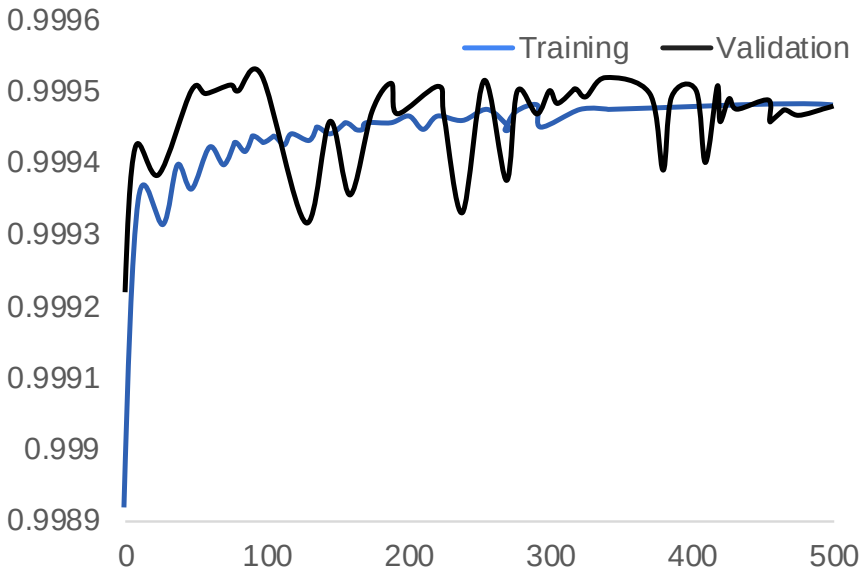

Fig. 4. Accuracy for 500 Iterations

In a duration of 20 s, an accuracy of 99% is obtained by logistic regression. When compared to RNN, LR obtained lesser accuracy despite the reduction in time duration. A categorical outcome may be observed from the LR. The accuracy of several machine learning algorithms may not be enhanced after a certain point at which they will not be able to handle bug data or predict new patterns. The new fraudulent patterns may be identified and the algorithm can adapt dynamically despite the complex data patterns when deep learning techniques are used. In the context of financial fraud detection, each machine learning technique has specific drawbacks. LR is vulnerable to over-fitting, the results of SVM are not transparent and RF can handle only small volumes of data.

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

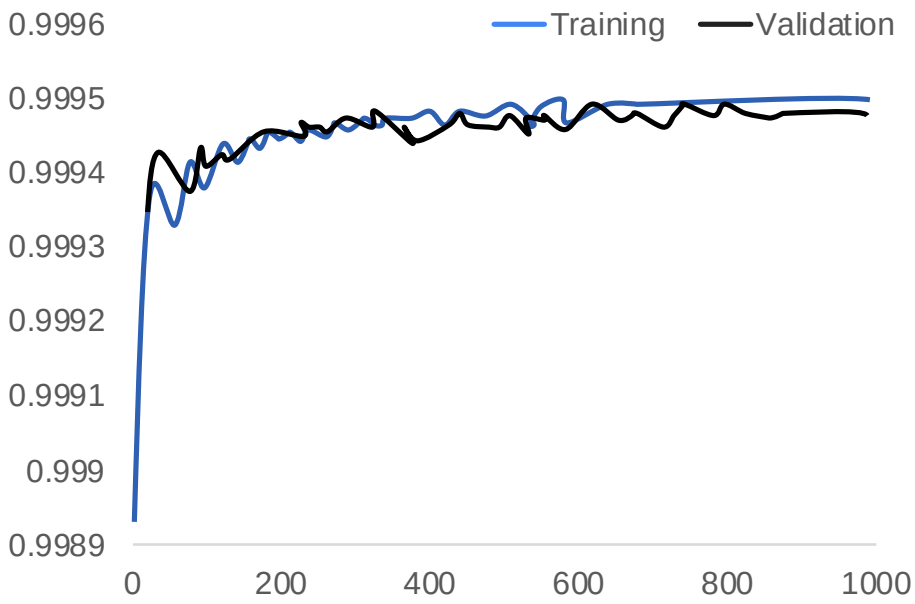

Fig. 5. Accuracy for 1000 Iterations

5. Conclusion

The financial sector and its stakeholders are affected with far-reaching implications due to financial frauds. Over the recent years, several challenges are compounded despite the emerging technological growth and reliance. In this age of big data, the traditional schemes remain inefficient. Real time credit card fraud dataset is used for DCNN based financial fraud detection in the proposed model. Despite the massive volume of data, the model offers improved detection, when compared to the existing models. The deep learning technique offers high accuracy and quick pattern identification for detecting the sophisticated and unknown patterns. This model addresses the inefficiency issues of the existing models. The existing machine learning models, auto-encoder models and other deep learning models are compared with the proposed model for performance evaluation by incorporating a real-time credit card fraud dataset. Fraud location, timing calculation and various other features may be incorporated with a single algorithm in future work.

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

References

[1] Zhang, R., Zheng, F., & Min, W. (2018). Sequential behavioral data processing using deep learning and the Markov transition field in online fraud detection. arXiv preprint arXiv:1808.05329.

[2] Joe, Mr C. Vijesh, and Jennifer S. Raj. "Location-based Orientation Context Dependent Recommender System for Users." Journal of trends in Computer Science and Smart technology (TCSST) 3, no. 01 (2021): 14-23.

[3] 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.

[4] Haoxiang, Wang, and S. Smys. "Overview of Configuring Adaptive Activation Functions for Deep Neural Networks-A Comparative Study." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 3, no. 01 (2021): 10-22.

[5] Choi, D., & Lee, K. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. Security and Communication Networks, 2018.

[6] Smys, S., and Jennifer S. Raj. "Analysis of Deep Learning Techniques for Early Detection of Depression on Social Media Network-A Comparative Study." Journal of trends in Computer Science and Smart technology (TCSST) 3, no. 01 (2021): 24-39.

[7] Wang, Y., & Xu, W. (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. Decision Support Systems, 105, 87-95.

[8] Ranganathan, G. "A Study to Find Facts Behind Preprocessing on Deep Learning Algorithms." Journal of Innovative Image Processing (JIIP) 3, no. 01 (2021): 66-74.

[9] Kim, E., Lee, J., Shin, H., Yang, H., Cho, S., Nam, S. K., ... & Kim, J. I. (2019). Champion-challenger analysis for credit card fraud detection: Hybrid ensemble and deep learning. Expert Systems with Applications, 128, 214-224.

[10] Vivekanadam, B. (2020). Analysis of Recent Trend and Applications in Block Chain Technology. Journal of ISMAC, 2(04), 200-206.

[11] Chakrabarty, Navoneel, and Sanket Biswas. "Navo Minority Over-sampling Technique (NMOTe): A Consistent Performance Booster on Imbalanced Datasets." Journal of Electronics 2, no. 02 (2020): 96-136.

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

[12] Błaszczyński, J., de Almeida Filho, A. T., Matuszyk, A., Szeląg, M., & Słowiński, R. (2021). Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. Expert Systems with Applications, 163, 113740.

[13] Hariharakrishnan, Jayaram, and N. Bhalaji. "Adaptability Analysis of 6LoWPAN and RPL for Healthcare applications of Internet-of-Things." Journal of ISMAC 3, no. 02 (2021): 69-81.

[14] Patil, V., & Lilhore, U. K. (2018). A survey on different data mining & machine learning methods for credit card fraud detection. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 3(5), 320-325.

[15] Shakya, Subarna, Lalitpur Nepal Pulchowk, and S. Smys. "Anomalies Detection in Fog Computing Architectures Using Deep Learning." Journal: Journal of Trends in Computer Science and Smart Technology March 2020, no. 1 (2020): 46-55.

[16] Al-Shabi, M. A. (2019). Credit card fraud detection using autoencoder model in unbalanced datasets. Journal of Advances in Mathematics and Computer Science, 1-16.

[17] Hamdan, Yasir Babiker. "Faultless Decision Making for False Information in Online: A Systematic Approach." Journal of Soft Computing Paradigm (JSCP) 2, no. 04 (2020): 226-235

[18] Singh, A., & Jain, A. (2020). An Empirical Study of AML Approach for Credit Card Fraud Detection–Financial Transactions. International Journal of Computers Communications & Control, 14(6), 670-690.

[19] Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. Applied Soft Computing, 106384.

[20] Adam, Edriss Eisa Babikir. "Survey on Medical Imaging of Electrical Impedance Tomography (EIT) by Variable Current Pattern Methods." Journal of ISMAC 3, no. 02 (2021): 82-95.

Authors Biography

Joy Iong-Zong Chen is currently a full professor in the Department of Communication Engineering Dayeh University at Changhua Taiwan. Prior to joining Dayeh University, he

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021

Journal of Artificial Intelligence and Capsule Networks (2021)
Vol.03/ No.02
Pages: 101-112
http://irojournals.com/aicn/
DOI: https://doi.org/10.36548/jaicn.2021.2.003

worked at the Control Data Company (Taiwan) as a technical manager from Sep. 1985 toSep. 1996. His research interests include wireless communications, spread spectrum technical, OFDM systems, and wireless sensor networks. He has published a large number of SCI Journal papers on the issues addressed by the physical layer for wireless communication systems. Moreover, he also majors in developing some applications of the IOT (Internet of Thing) techniques and Dr. Joy I.-Z. Chen owned some patents authorized by the Taiwan Intellectual Property Office (TIPO).

Kong-Long Lai, is currently working as professor in the Department of Electrical Engineering, Da-Yeh University, No. 168 University Rd., Dacun, Changhua 51591, Taiwan. His area of research includes Information and communication technologies, Antenna design and RF propagation, Autonomic, & Dependable computing, Green Computing, Computational Complexity and Networks

ISSN: 2582-2012 (online)
Submitted: 8.05.2021
Revised: 29.05.2021
Accepted: 15.06.2021
Published: 22.06.2021