*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.7850*
***Article Received**: 28 April 2023 **Revised**: 22 June 2023 **Accepted**: 05 July 2023*

# A Novel Method of Fraud Detection of Credit Cards by Fuzzy, LSTM, and PSO Optimization

**P. Kavya Sree[1], V. Kakulapati[2], B. Indira[3]**
[1]Sreenidhi Institute of Science and Technology
Yamnampet, Ghatkesar, Telangana-501301, India
kavyasripapani@gmail.com
[2]Sreenidhi Institute of Science and Technology
Yamnampet, Ghatkesar, Telangana-501301, India
vldms@yahoo.com
[3]Sreenidhi Institute of Science and Technology
Yamnampet, Ghatkesar, Telangana-501301, India.
bindira@sreenidhi.edu.in

**Abstract:** Since online shopping has become so popular, credit card theft has skyrocketed. This makes it hard to spot fake charges on accounts. In this research, credit card fraud detection is performed using a fuzzy database. It has been considered a data mining challenge to reliably identify whether or not a transaction is legitimate. This paper discusses the LSTM method and fuzzy logic. The learning process was sped up and made more precise by using a technique called particle swarm optimization (PSO). Performance values have been compared with those of the LSTM and fuzzy logic techniques, and a PSO-based neural network has been intensively trained by increasing the number of iterations and the population, or number of swarms. It has been shown that the PSO-based algorithm yields the best result for detecting fraudulent transactions. The goal of this method is to lower the mean square error (MSE) rate of the system. PSO is a popular optimization technique that can be used to locate the optimal set of features for the credit card fraud detection system. The proposed method PSO shows less mean squared error compared with Fuzzy and LSTM techniques.

***Keywords:*** Credit Card, Fraud detection, fuzzy logic, LSTM, PSO. ML, methods, optimization, MSE, optimal.

## I. INTRODUCTION

One of the most reported types of fraud is committed using credit cards. It is possible to avoid credit card theft by using authorization techniques such as signatures, credit card numbers, identification values, the cardholder's mailing address, the expiration date, etc., but these procedures are not infallible. Approaches to fraud detection for credit card fraud need to examine data that can identify and eradicate fraud. These days, online purchases account for 50 percent of all credit card fraud, and simple example matching algorithms aren't enough to identify fraudulent activities. Success in uncovering fraud should confirm the precision and low volume of false positives.

The scientific community is interested in credit card fraud because it is hard to find. One of them is the fact that there are many more legitimate transactions than fraudulent ones in credit card fraud data sets. Traditional classifiers frequently use specific transaction information, such as amount, timing, and location, to determine whether a transaction is likely fraudulent or not [1]. Consumer purchasing patterns, which help uncover related fraud trends that change over time owing to regularity and new attack methodologies [2], are not taken into account by these methods.

Due to the large quantity of data that is now accessible to organizations and the increase in computing capability, ML approaches have become more potent and affordable for solving increasingly challenging issues in the modern world. To identify fraudulent applications and transactional deception, several ML techniques are applied. Machine learning methods approximate data extraction by instructing a machine to accomplish a difficult task. To analyze data with little processing, the adaptive nature of deep learning methods, a subset of ML, is generally advantageous. Assigning the task of feature engineering to a machine saves time and allows non-experts to participate in the analysis phase.

The use of fuzzy logic is an approach to processing that makes allowances for imprecision and erroneous data. This is a crisp rules method in which the certainty variables are integers between 0 and 1. Perceptions may be true (1.0) or false (0.0) and the reasoning can be used anyway, even if the perception is utterly off (0.0)]. In order to optimize the MLP training process, the particle swarm optimization (PSO)

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.7850*
***Article Received**: 28 April 2023 **Revised**: 22 June 2023 **Accepted**: 05 July 2023*

method is employed to find the best possible weights and biases. It has been shown that the PSO algorithm outperforms other heuristic optimization techniques. PSO is easy to construct, fast on the CPU, and forgiving concerning the parameters it uses to make decisions.

The remaining sections of the work are structured into the following categories: Fraudulent use of credit cards is discussed in detail in Sec. 2. In Sec. 3, we discussed an approach to addressing fraud detection via analysis. In Sec. 4, describe the recommended imprecise, LSTM, and PSO optimizations. Sec. 5 analyses the method of experimentation and outcomes. The concluding remarks are discussed in sec 6 followed by future enhancement.

## II. LITERATURE SURVEY

A fuzzy hereditary programming framework [4] to solve the problem of discovering patterns in massive data sets. This is accomplished by using a collection of fuzzy recommendations of varying lengths to be summarized using a set of designated classes. In addition to introducing the cycles of modular development, modular re-development, and settled transformative pursuit, we also provide some additional techniques, such as the use of genotypes in hereditary programming and two novel hybrid hereditary administrators. Evidence from experiments shows that the framework successfully classifies data from the Wisconsin Bosom Malignant Growth Data Set 95% of the time.

Top-to-bottom measurable examination is regularly required considering information normalization and utilization in counterfeit brain organizations. This study looks at the utilization of bunch examination considering information standardization in a situation study that includes the distinguishing proof of visa extortion. Neuronal sources of information can be limited through bunching properties, as indicated by starter discoveries from the utilization of fake brain organizations and group investigation to the identification of misrepresentation. [5].

One critical utilization of expectation calculations is the decrease in Visa extortion. The high required symptomatic quality is a critical obstruction to the utilization of the brain network in preparing strategies. Since only one monetary exchange in 1,000 is invalid, any expectation of accomplishment below 99.9% is unsuitable. These MasterCard exchange prerequisites required the creation and testing of totally different ideas using genuine Visa information. This study demonstrates how state-of-the-art information mining techniques and brain network calculation can be effectively coordinated to achieve superb misrepresentation inclusion and a low deception rate. [6].

In this work, we examine AI-based computerized Visa extortion recognition. Visa extortion discovery is essential considering monetary associations in this advanced age. On genuine monetary information, we utilize two AI procedures—counterfeit brain organizations and Bayesian conviction organizations—appropriate considering thinking under vulnerability to the main thing in need of attention. To wrap things up, new bearings are recommended to improve the two techniques and results. [7].

Because of how quickly e-commerce has grown, more people are using charge cards. There have been more cases of visa extortion, which is a form of financial fraud that costs the government billions of dollars every year. To fix this problem, all banks that give out credit cards should put in place strong procedures for recognizing misrepresentation right away. Different types of credit card fraud can now be distinguished using cutting-edge techniques [8].

A comparison of the effectiveness of ANN, RF, and SVM approaches with CNN and LSTM deep learning approaches [9] designed for identifying credit card fraud. The suggested approaches fared better in experiments than conventional ML models in identifying credit card fraud. To reorganize actual transaction characteristics into discrete convolutional features for the anticipated [10] deep learning-based fraud detection strategy for transactions made online. As assessed against the current CNN for detecting fraud, it can stabilize both recall and accuracy at around 91% and 94%, accordingly, leading to a gain of 26% and 2%, respectively.

## III. METHODOLOGY

Fraud detection of Credit Cards can be detected by many traditional means such as; Skimming, Phishing, Chargeback cards, Hacking, and Identity Theft. All these come in the form of fraud, which means for causing Credit Card fraud it is possible to classify credit card fraud into three categories they are Legal, Suspicious, and Fraud. This strategy is based on the credit card number, Active transactions, and time interval in which the credit card is used.

Fraud detection is a critical issue in the banking and finance sector. With the advancement in technology, fraudulent activities have become more sophisticated, and traditional fraud detection methods are no longer sufficient. Therefore, there is a need for more advanced and intelligent techniques to detect fraudulent activities [11].

Fuzzy logic and LSTM (Long Short-Term Memory) are two popular techniques used in credit card fraud detection. Fuzzy logic is a mathematical technique that deals with uncertainty and imprecision in data. It is particularly useful in fraud detection, as fraudulent activities can often be

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.7850*
***Article Received**: 28 April 2023 **Revised**: 22 June 2023 **Accepted**: 05 July 2023*

ambiguous and difficult to detect using traditional methods. Fuzzy logic can be used to create rules that can identify suspicious activities based on a set of predefined criteria [12]. This work analyzed the fuzzy model to estimate the accuracy of fraud detection of credit cards. Expected values are assessed with experimental values for precision to be produced.

The vanishing gradient issue may be addressed and long-term dependencies in sequential data can be captured using a recurrent neural network (RNN) architecture called Long Short-Term Memory (LSTM). The LSTM field is flourishing and undergoing extensive study. It is not easy to understand how LSTMs work in a bidirectional setting or how line-to-sequence ideas apply here. The RNN Special Model with LSTM prediction was presented in [13]. To carry out LSTM, the model RNN must exploit the information in a previously buried n-level layer. Measuring and computation become progressively more time-consuming and resource-intensive because of this finding. For both long-term and short-term computations, RNN methods are difficult to utilize using the same underlying concepts.

Data classification and regression using the algorithm for Random Forests, an approach for supervised learning. During the model training phase of the classification operation, the method builds several decision trees, each of which is used in the construction of the classification and the subsequent technique of class yield [14].

PSO (Particle Swarm Optimization) is a computational approach for the calculation of population size and one of the fundamental optimizations used. To go closer to the nearest partner at any given time, the values of a set of variables are adjusted numerous times. To get the best possible optimum approach this method is employed as the operation. Only through evaluating a function can information be gleaned. The procedure is simple to carry out. It seems that the PSO remains able to construct probable particles and explore for an optimum response [15].

## IV. IMPLEMENTAITON ANALYSIS

To create more useful autonomous systems, this work aims to improve & optimize existing processes. The application of fuzzy logic is the main topic of this work. The usual set theory says that the output can be either zero or one. While the representation of ambiguous & uncertain information through logical operators led to the development of fuzzy logic. People frequently think in uncertain ways, therefore fuzzy logic's rules are simple even considering individuals who are not familiar with the idea. To accurately identify fraudulent, suspect, or legitimate credit card transactions, the author of this research applies Fuzzy Logic membership functions along with PRO-LSTM Methods.

**1.1.** *Description of Data***:** The data set was gathered from various sources (Kaggle, GitHub, etc.). To ensure the accuracy of the study's results many unnecessary columns are eliminated. The dataset obtained after many corrections consist of **7** columns and **29001** rows. The first columns mention the Credit Card Number and the remaining columns mention the prediction factors.

After collecting data, Pre-Processing for cleaning data for reducing missing and noise elements. Generate a response variable in a categorical form which takes values 1 and 0:

1= counterfeit activity

0 = normal activity

**1.2.** *Model Building and Evaluation***:** Using Fuzzy Logic, Long Short-Term Memory Networks (LSTM), and Random Forest to split the data and train the data and this process can be done only after labeling the classes for prediction to avoid complexity. The last column of the data is used for testing and the remaining 6 subsets are used for training. The evaluation of the models used in this study is measured through the performance metrics and Mean Square Error (MSE) of the models.

Fuzzy logic is a type of logic that deals with reasoning that is approximate rather than precise. It is based on the idea that things can have degrees of membership in a set, rather than being either completely in or completely out of the set.

In traditional "crisp" logic, a proposition is either true or false, but in fuzzy logic, a proposition can have a truth value between 0 and 1, representing the degree to which it is true. For example, the statement "it is hot outside" might have a truth value of 0.8, indicating that it is mostly true but not completely true.

| | Unnamed: 0 | trans_date_trans_time | cc_num | ... | merch_lat | merch_long | is_fraud |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 2019-01-01 00:00:18 | 2703186189652095 | ... | 36.011293 | -82.048315 | 0 |
| 1 | 1 | 2019-01-01 00:00:44 | 630423337322 | ... | 49.159047 | -118.186462 | 0 |
| 2 | 2 | 2019-01-01 00:00:51 | 38859492057661 | ... | 43.150704 | -112.154481 | 0 |
| 3 | 3 | 2019-01-01 00:01:16 | 3534093764340240 | ... | 47.034331 | -112.561071 | 0 |
| 4 | 4 | 2019-01-01 00:03:06 | 375534208663984 | ... | 38.674999 | -78.632459 | 0 |

[5 rows x 23 columns]

Fig 1: Evaluation of Fuzzy membership function of uploaded data

In this work, consider 0 & 1 as MEDIUM & 2 as HIGH

| | Card_No | Time_Difference | Amount_Difference | Location | Interval | Frequency | label |
|---|---|---|---|---|---|---|---|
| 0 | 60422928733 | 2 | 2 | 2 | 1 | 0 | Fraud |
| 1 | 60422928733 | 2 | 2 | 2 | 1 | 0 | Fraud |
| 2 | 60422928733 | 2 | 2 | 2 | 1 | 0 | Fraud |
| 3 | 60427851591 | 2 | 2 | 2 | 1 | 0 | Fraud |
| 4 | 60487002085 | 2 | 2 | 2 | 1 | 0 | Fraud |

Fig 2: Detecting Fraud transactions using Fuzzification

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.7850*
***Article Received**: 28 April 2023 **Revised**: 22 June 2023 **Accepted**: 05 July 2023*

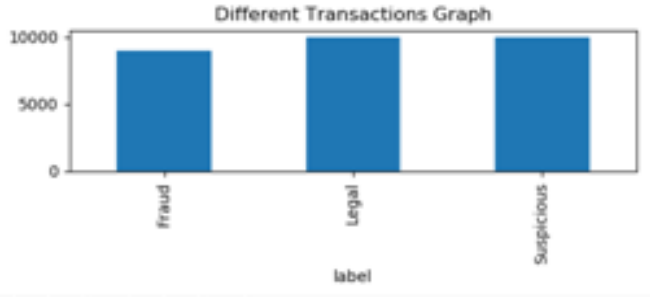


Fig 3: Type of transaction Vs number of records

Table 1. Values of the metrics.

| Name of the metric | Measure of the metric |
|---|---|
| Accuracy | *(TP+TN)/((FP+FN+TP+TN))* |
| Precision | *TP/((FP+TP))* |
| Recall | *TP/((FN+TP))* |

The letters TP, TN, FP, and FN thus stand for true positive, true negative, false positive, and afterward false negative. The dataset has n data points, y_true represents the target variable's true value, and y_pred represents the model's anticipated value.

The four prediction models that were tested in this book to determine the amount of student involvement and which student activities were more crucial are the most crucial pieces of information. The findings demonstrated that credit Card Fraud may be anticipated by ML algorithms, with the LSTM method having high precision and an extremely low MSE. The fuzzy logic method was utilized to identify data unpredictability, and the random forest technique was utilized to identify key traits that may be used to forecast instances of fraud using credit cards. The correctness of the LSTM method was outstanding, and the MSE of the algorithm used for it was extremely small, demonstrating that credit cards may be detected using machine learning techniques.

Table 2. Performance of the Models

| Model Name | Accuracy | Precision | F-1 Score | Recall |
|---|---|---|---|---|
| Fuzzy Logic | 0.96 | 0.94 | 0.91 | 0.9 |
| PSO-LSTM | 0.98 | 0.96 | 0.94 | 0.94 |
| Random Forest | 0.96 | 0.93 | 0.9 | 0.9 |

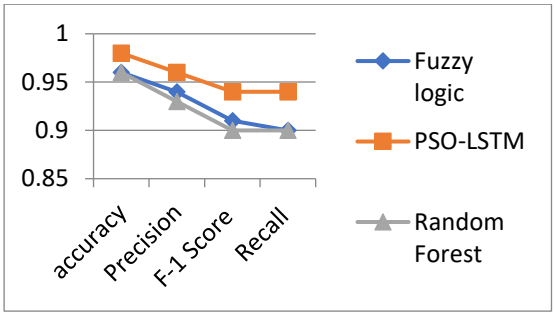


Fig 4: The proposed model performance measures.

Mean squared error (MSE) is a popular statistic for assessing the success of an ML regression problem. The mean squared error among a dataset's expected and actual values is calculated. By averaging the squared discrepancies between the anticipated outcomes and the actual results, we get the MSE, with a smaller number indicating better effectiveness of the model.

PSO optimized features algorithm which will select optimal features from the dataset and these optimized features will get retrained with LSTM to reduce the MSE error rate. The lower the MSE the better is the algorithm. In the below screen, we are showing code for PSO features optimization

The below table shows the MSE values of the models of this study.

Table 3. MSE Values of the classification models.

| *Model Name* | *MSE Values* |
|---|---|
| *Fuzzy Logic* | *0.82* |
| *PRO-LSTM* | *0.02* |
| *Random Forest Classifier* | *0.1* |

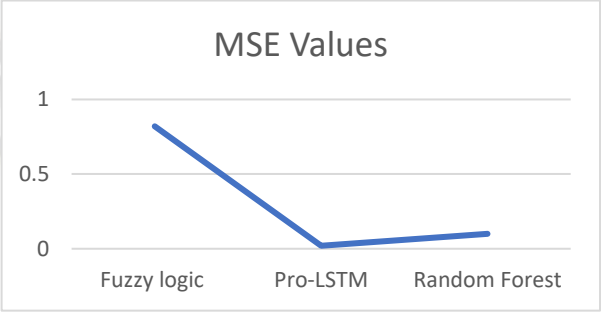


Fig 5: MSE Values of the classification models.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.7850*
*Article Received: 28 April 2023 Revised: 22 June 2023 Accepted: 05 July 2023*

Table 4. Salient features of data

| Feature number | Feature Name | Feature Importance |
|---|---|---|
| *1* | *Card_No* | *0.17* |
| *2* | *Time_Difference* | *0.24* |
| *3* | *Amount_Difference* | *0.47* |
| *4* | *Location* | *0.025* |
| *5* | *Interval* | *0.096* |
| *6* | *Frequency* | *0.0* |

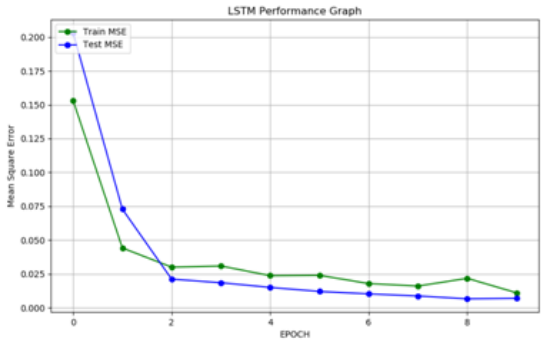


Fig 6: The performance of the LSTM network using Mean Square Error

Among each increasing epoch MSE got decreased & considering any model decreasing MSE consider as best mode.

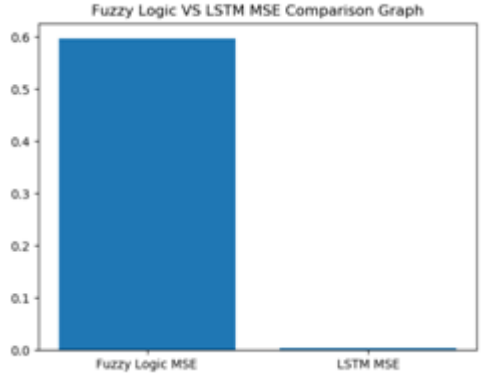


Fig 7: The comparison between fuzzy logic MSE and LSTM MSE

Fig 5 exhibits LSTM MSE is very low when compare to Fuzzy which implies that LSTM is better than the Fuzzy algorithm

*PSO*: This is an optimized features algorithm that selects optimal features from the dataset & these optimized features will get retrained among LSTM to reduce the MSE error rate. The lower the MSE the better is the algorithm.

FUZZY LOGIC MSE performance is 0.87% and LSTM MSE is 0.0027%.

Then applied optimization algorithm and the performance of PSO LSTM MSE is 0.0017 which is lesser than normal LSTM.

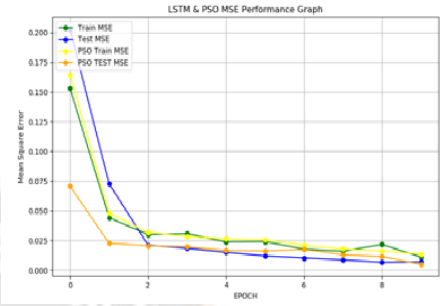


Fig 6: The performance of the LSTM network and PSO optimization using Mean Square Error

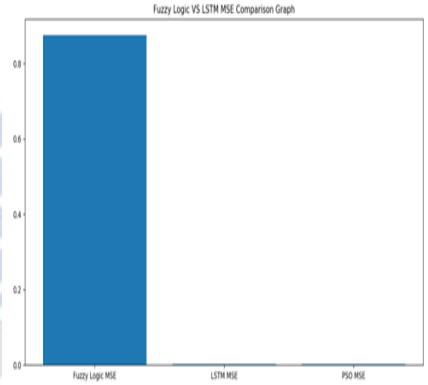


Fig 7: The comparison between fuzzy logic MSE and LSTM MSE

In the figure, PSO MSE error is too small when compared to FUZZY PSO can reduce less MSE error rate.

## V. CONCLUSION

To reduce the computational time required to identify fraudulent transactions & to minimize false alarms, credit card fraud detection techniques must be precise & quick. In our study, we employed a behavior credit card model to monitor how frequently people used their credit cards. The membership functions considering each attribute were generated using fuzzy logic after the data had been preprocessed to be divided into useable attributes. Based on the importance of each input, the rules were also developed & prioritized. To test & validate our findings, a deep learning LSTM model was deployed. Our findings showed that the LSTM model was more precise than the fuzzy logic model. When comparing the mean square error, the outcomes of the fuzzy logic are still regarded as acceptable. Our approach has the advantage of using the transaction statement's pre-existing data considering decision-making, which eliminates the requirement of considering specialized tools. As an addition,

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.7850*
***Article Received**: 28 April 2023 **Revised**: 22 June 2023 **Accepted**: 05 July 2023*

we have included the PSO optimization algorithm. When compared to other algorithms, this approach has a low MSE rate. In our upcoming work, we intend to implement our project using actual credit card transactions from many users & expand the scope of our research to include big data analysis that reflects the large volume of transactions that take place daily.

## VI. FUTURE ENHANCEMENT

In the future, the system can be enhanced by incorporating additional data sources, such as customer profiles, geographical location, and purchase history. These additional data sources can help improve the accuracy of fraud detection by providing more contextual information about the transaction. Exploring other deep learning techniques: While LSTM is a powerful technique for detecting patterns in sequential data, Various methods of deep learning, such as CNNs and Generative Adversarial Networks, may be investigated to improve the method's effectiveness. The proposed system can be enhanced by implementing real-time monitoring capabilities, allowing the system to detect and respond to fraudulent activities in real time.

## REFERENCES

[1] Donato JM, Schryver JC, Hinkel GC, Schmoyer RL, Leuze MR, Grandy NW. Mining multi-dimensional data for decision support. Future Gener Comput Syst. 1999;15:433–41.

[2] Dal Pozzolo A, Boracchi G, Caelen O, Alippi C, Bontempi G. Credit card fraud detection: realistic modeling and a novel learning strategy. IEEE Trans Neural Netw Learn Syst. 2018;29(8):3784–97.

[3] V. Kakulapati, et al., (2021). Fuzzy-Based Predictive Analytics for Early Detection of Disease—A Machine Learning Approach. In: Tuba, M., Akashe, S., Joshi, A. (eds) ICT Systems and Sustainability. Advances in Intelligent Systems and Computing, vol 1270. Springer, Singapore. https://doi.org/10.1007/978-981-15-8289-9_9

[4] Mallinson, H. et al., "Evolving Fuzzy Rules considering Pattern Classification" in International Conference on Computational Intelligence considering Modelling, Control & Automation - CIMCA'99.Vol. 1, IOS Press,1999, pp. 17- 19

[5] Mineda Carneiro, E.; Vieira Dias, L.A.; Da Cunha, A.M.; Stege Mialaret, L.F., "Cluster Analysis & Artificial Neural Networks: A Case Study in Credit Card Fraud Detection," in Information Technology - New Generations (ITNG), 2015 12th International Conference on, April 2015, pp.122-126, 13-15

[6] Tao Guo; Gui-Yang Li, "Neural data mining considering credit card fraud detection," in Machine Learning & Cybernetics, 2008 International Conference on, vol.7, July 2008, pp.3630-3634, 12-15.

[7] S. Maes, K. Tuyls, B. Vanschoenwinkel & B. Manderick, "Credit card fraud detection using Bayesian & neural networks," in Proceedings of the First International NAISO Congress on Neuro Fuzzy Technologies, 2002

[8] Tripathi K.K. & Pavaskar M.A. "Survey on Credit Card Fraud Detection Methods" in International Journal of Emerging Technology & Advanced Engineering, 2(11), November 2012.

[9] T. T. Nguyen, H. Tahir, M. Abdelrazek, and A. Babar, "Deep learning methods for credit card fraud detection," 2020 Dec, https://arxiv.org/abs/2012.03754.

[10] Z. Zhang, X. Zhou, X. Zhang, L. Wang, and P. Wang, "A model based on convolutional neural network for online transaction fraud detection," *Security and Communication Networks*, vol. 2018, Article ID 5680264, 2018.

[11] Asma Cherif, et al., Credit card fraud detection in the era of disruptive technologies: A systematic review, Journal of King Saud University - Computer and Information Sciences, Volume 35, Issue 1, 2023, Pages 145-174, ISSN 1319-1578,https://doi.org/10.1016/j.jksuci.2022.11.008

[12] Malini, N. & Pushpa, M. (2017). Analysis of credit card fraud identification techniques based on KNN and outlier detection. 255-258. 10.1109/AEEICB.2017.7972424.

[13] V. Kakulapati et al., "A Novel Multimodal risk disease prediction of Covid-19 by Using Hierarchical LSTM Methods ", "Taylor and Francis book " Data Science and Data Analytics: Opportunities and Challenges". ISBN 9780367628826

[14] V. Kakulapati et al., Risk analysis of coronaviruses caused death by the probability of patients suffering from chronic diseases - a machine learning perspective", International Journal Critical Reviews. 2020; 7(14): 2626-2633. Doi:10.31838/jcr.07.14.499.

[15] V.Kakulapati et al., "An Intelligent methodology for covid 19 risk prediction using swarms intelligence optimization- A machine learning perspective" Taylor and Francis book "Swarm Intelligence and Machine Learning: Applications in Healthcare". SIMLAH-2021, ISBN 9781032145792, DOI: 10.1201_9781003240037-1, PP: 1-20. Sep 29, 2022