(IJAER) 2021, Vol. No. 22, Issue No. I, July **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

LSTM NETWORK AND WOA DRIVEN INTELLIGENT METHOD FOR LOAN ELIGIBILITY PREDICTION

Agnalin V

Student of Department of Computer Science and Engineering, Bethlahem Institute of Engineering, Karungal

ABSTRACT

People working in banks face many problems, including approving loans. The project proposes a loan sanctions system that determines whether a loan should be provided to someone based on certain attributes. Although the bank follows strict rules and regulations, and conducts detailed background checks when approving loans, and keeps in mind the individual's ability to repay the loan, in this case, people often fail to repay the loan and have already given it to him. In this project, the system , propose for bankers will help them predict the trusted customers who have applied for loans, thereby increasing the chance of timely repayment of loans. This prediction is done using LSTM neural network. To improve the performance of LSTM , Whale optimization algorithm is used in this project.

Keywords:- Loan eligibility, Classification, LSTM, WOA, optimized.

1. INTRODUCTION

Machine learning (ML) is a subfield of artificial intelligence (AI) [18]. The goal of machine learning is usually to understand the structure of the data and fit the data to a model that people can understand and use. Although machine learning is a field of computer science, it is different from traditional computing methods. In traditional computing, an algorithm is a set of clearly programmed instructions that a computer uses to calculate or solve problems. In contrast, machine learning algorithms allow computers to train data input and use statistical analysis in order to output values that fall within a specific range. Therefore, machine learning helps computers build models based on sample data, thereby automatically executing decision-making processes based on data input.

Machine Learning has become one of the mainstays of information technology and with that, a rather central, albeit usually hidden, part of our life. Machine learning can appear in many guises [7]. Machine Learning is the field of study that gives computers the capability to learn without being explicitly programmed. ML is one of the most exciting technologies that one would have ever come across. Machine learning is the technology that allows systems to learn directly from examples, data, and experience [35].

Machine learning is used to teach machines how to handle the data more efficiently. Sometimes after viewing the data, we cannot interpret the pattern or extract information from the data. In that case, we apply machine learning. With the abundance of datasets available, the demand for machine learning is in rise. Many industries from medicine to military apply machine learning to extract relevant information. The purpose of machine learning is to learn from the data. Many studies have been done on how to make machines learn by themselves [4]. Loan forecasts are very helpful for bank employees and applicants. The purpose of Machine learning is to provide a quick, direct and easy way to select qualified applicants. It can provide banks with special advantages. ML is used to predict the candidate eligible or not for the loan. The property of good loan eligible prediction system are: a) It must be able to identify the not eligible person accurately, b) It must quickly predict the eligible or not eligible cases, c) At any case a eligible person should not be consider as a not eligible person.

Machine learning algorithms have the ability to improve themselves through training. Today, three prominent methods are used to train ML algorithms. These are the three types of machine learning: supervised learning, unsupervised learning and reinforcement learning

2. REVIEW OF RELATED WORKS

Due to the massive rise [10] in the development in the economy, there has been huge increment in the requirements of customer s personal loans. As the borrower's behaviors are too constant, have high uncertainty and are fuzzy in nature. The banks are very much anxious about the repayment of the money. In this study, the authors dig the default patterns from the big data of the banks. The authors used the SOM for dividing the borrowers into various groups. Finally, the behavior and the characteristics of the borrowers of every group are analyzed, so that the credit risk can be lowered for the consumer loans.

This article [53] examines the four types of the Chinese commercial banks for examining the cost and profit efficiency from 2002- 2013. They found that the domestic banks are less cost efficient and more profit efficient whereas the foreign banks are least profit efficient and having most cost efficient. After WTO transition period, the profit efficiency gap has widened between the domestic and foreign banks. The authors also used Auto Regression method to find a relationship within the efficiency and SROE.

The authors [12] used the data of the commercial and deposit banks to determine and estimate the attributes for the banks risk. According to investors, the banks having small loan maturity positions are safer for investment which increases the trading income.

In banking sector retails loans play a key role in many countries. Loans provided to individuals are regarded as riskier then loans given for business person. The author in this article [36] used the fuzzy logic model for the loan evaluation. In this model, it consists of five input variables: income, employment, character, credit history which shows the standing credit. Based on the degree of the membership of the linguistic terms of the fuzzy output, the applicants standing credit should be classified as high, medium, low.

In this article [6] the author proposed an algorithm that is SCAD- Simultaneous Clustering and Attribute Discrimination which performs feature weighting and clustering simultaneously. This algorithm is used for the analyses of the repaid information of the bank loan, which can efficiently find the associated weight of the loan information factors and can realize the potential customer.

In this article [49], the author applied machine learning Technology and preprocessing algorithms to analyze Credit risk involved in peer-to-peer loans system. These techniques are used to explore, analyze And determine the factors that play an important role Used to predict the credit risk in Lending Club.

In this article [52], the author measured Bank loans based on prospect theory. Company Profile People with lower credit ratings have difficulty getting loans from banks bank. The author found the direct or Indirectly affect the willingness and performance of the loan Some certain measures, in return, these measures are used to solve The company's financial issues.

In this article [1], the author uses data mining techniques For example, the K-mean method, which is clustering technology and decision tree, is used to analyze customers' personal loans and Used to analyze the customer's payment performance Reduce the risk rate in decision-making.

In this article [56], the author collected bank loan data There are 96 features and 10415 samples. Author Data mining methods such as relief algorithms and data mining are used PCA is used to analyze and extract the characteristics of non-performing loans in NPL Commercial banks through relatively bad and Perform loan records. This research can find non-performing loans, Capture warning signs from commercial banks.

Author [57] used for rough set and BP neural network Internet to find the default risk of personal loans. they First construct the default index of personal loans, and then Used for streamlined rough set, and then used on sample BP The neural network is trained.

Based on probability of default and risk capital RAROC The method integrates the pricing method. The method [26] is Used to analyze the pricing status of bank loans. analysis Various attributes, such as credit risk, identifying risk types, The probability distribution of the loan.

In this article, the author [40] uses such as Logistic regression and CART and such as The loss matrix, matrix weights and prior probabilities Manage unbalanced data. The author focuses on Credit risk in bank data using machine learning technology.

In this article, the author [9] analyzes the risk. The Finnish Interbank Payment System is the most used The entropy idea of estimating transaction matrix payment Then show the mechanism of risk.

(IJAER) 2021, Vol. No. 22, Issue No. I, July **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

In this article [54], the author uses Bayesian rule, so The bank can correctly determine the type of business in order to Avoid making mistakes when determining loans. Overall In the process, the author uses game theory to find Issues related to inter-enterprise credit financing And the bank.

In this article [55], the author applied SVM or implemented Analysis of corporate debt repayment data Reduce the risk of banks providing loans. they Compare the classification results of Logistic regression And neural network evaluation model. The main points of the paper SVM-based classification is better than Logistic Return after comparison.

3. PROPOSED LSTM NETWORK AND WOA DRIVEN INTELLIGENT METHOD FOR LOAN ELIGIBILITY PREDICTION

This section This section describes the peculiarities and operating conditions of the loan eligibility system. The proposed method illustrates Three layers of control such as, 1) Pre-processing of data 2) Feature selection, 3)Long-Short Time Memory(LSTM) network and WOA. Figure 1 shows the block diagram of the proposed LSTM network driven intelligent method for loan eligibility prediction.

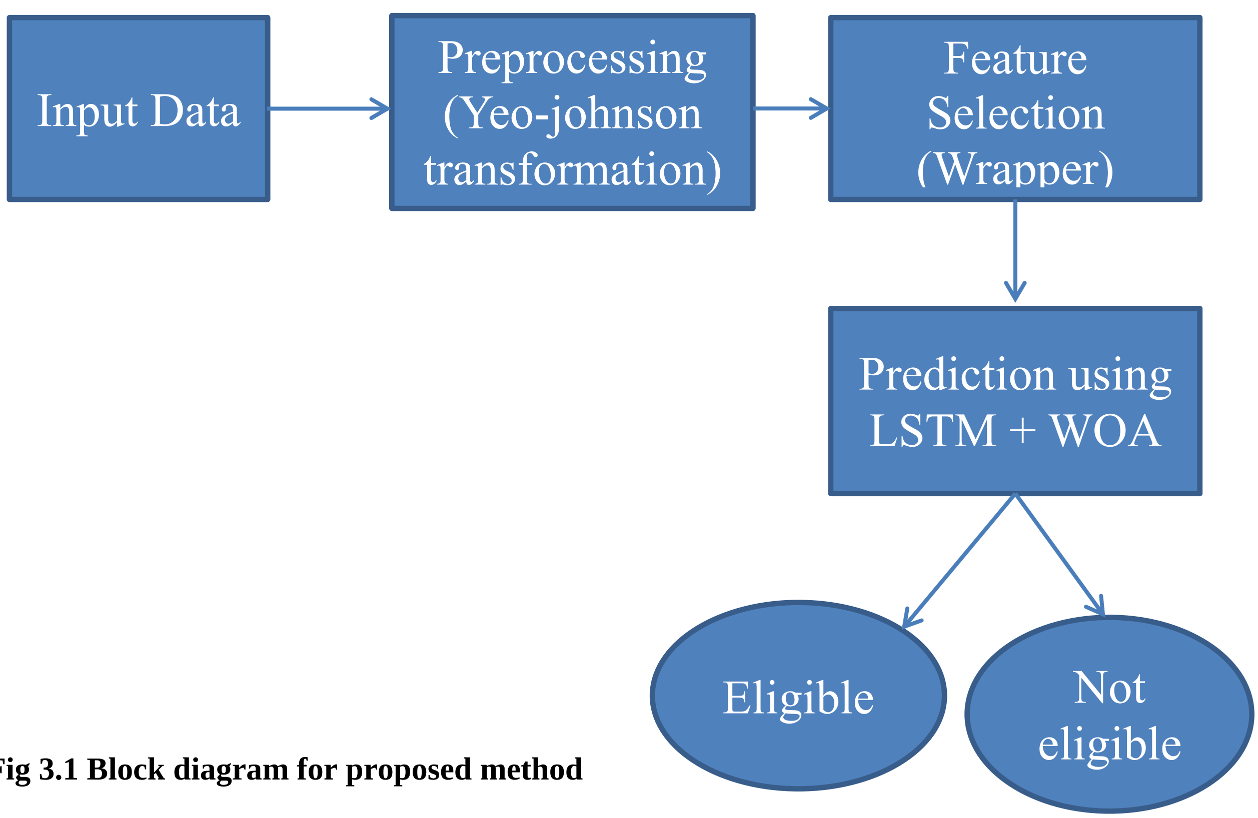


Fig 3.1 Block diagram for proposed method

3.1.1 Pre-processing

Feature extraction is one of the most important pre-processing steps [7]. Data pre-processing is the process of preparing raw data and making it suitable for machine learning models. This is the first step in creating a machine learning model, and it is also a crucial step. When creating machine learning projects, it's not always the case with neat and formatted data. Moreover, when performing any operations on the data, it must be cleaned up and placed in a formatted manner.

Actual data usually contains noise, missing values, and may be in an unusable format that cannot be directly used in machine learning models. Data pre-processing is a necessary task to clean up data and make it suitable for machine learning models, which also improves the accuracy and efficiency of machine learning models. Data pre-processing is a necessary step before building a model with these functions.

3.1.2 Yeo-johnson transformation

The Yeo-Johnson transformation is an extension of the Box-Cox transformation and can be used for variables with zero and negative and positive values. This transformation is typically done on the outcome variable using the residuals for a statistical model (such as ordinary least squares). Here, a simple null model (intercept only) is used to apply the transformation to the predictor variables individually. This can have the effect of making the variable distributions more symmetric.

3.2 FEATURE SELECTION

Feature selection is one of the core concepts in machine learning, which greatly affects the performance of the model. The data function used to train the machine learning model has a great influence on the achievable performance. Feature selection and data cleaning should be the first and most important step of model design.

Feature selection is the process by which you automatically or manually select those features that contribute the most to the predictor variable or output. Having irrelevant features in the data will reduce the accuracy of the model and make the model learn based on irrelevant features.

3.2.1 Wrapper Method

It select particular number of feature from pre-processing data. The working principle of the packaging method is to use a machine learning algorithm to evaluate feature subsets. The algorithm uses a search strategy to browse the space of possible feature subsets, and then evaluates each subset according to the performance of a given algorithm.

These methods are called greedy algorithms because their purpose is to find the best possible combination of features that lead to the best performance model, which will be computationally expensive and often impractical in the case of exhaustive searches. In fact, any combination of search strategies and machine learning algorithms can be used as a wrapper.

3.2.1.1 Advantage

- They detect the interaction between variables
- They find the optimal feature subset for the desired machine learning algorithm

3.2.1.2 Exhaustive Feature Selection

This method tries all possible feature combinations. This is the most robust feature selection method covered so far. This is a brute-force evaluation of each feature subset. This means that it tries every possible combination of the variables and returns the best performing subset.

3.3 LSTM NETWORK

Once the features are selected, the feature vector is given to LSTM network layer. LSTM network predict the loan eligible candidate. An LSTM has a similar control flow as a recurrent neural network (RNN). It processes data passing on information as it propagates forward. The differences are the operations within the LSTM's cells. LSTM created as the solution to short-term memory. They have internal mechanisms called gates that can regulate the flow of information.

These gates can learn which data in a sequence is important to keep or throw away. By doing that, it can pass relevant information down the long chain of sequences to make predictions. Almost all state of the art results based on recurrent neural networks are achieved with these two networks. These operations are used to allow the LSTM to keep or forget information.

3.3.1 Architecture

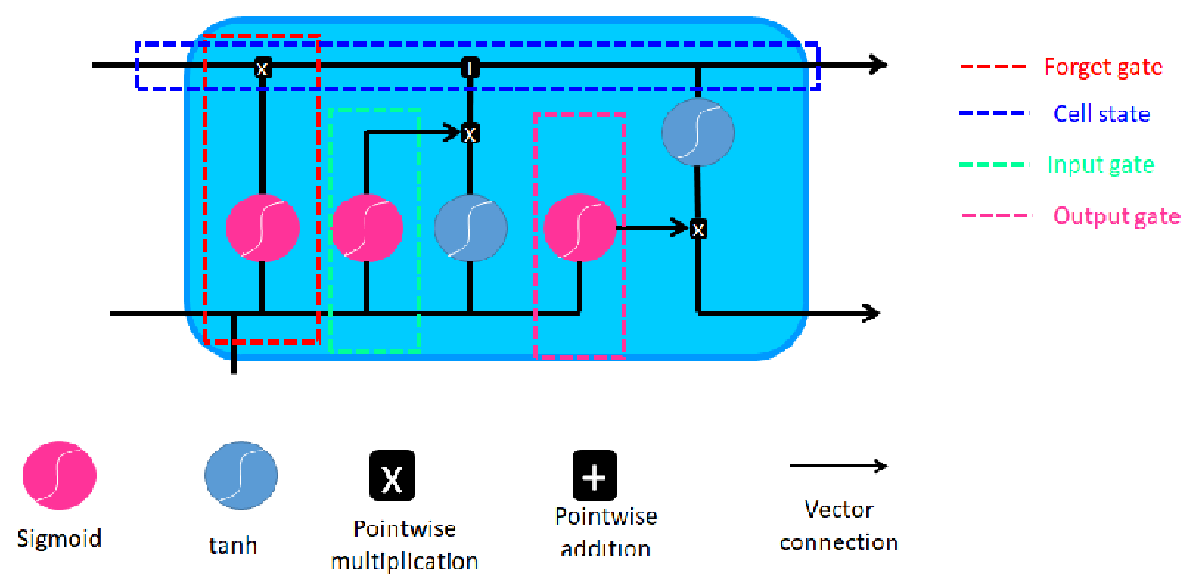


Fig 3.2 Architecture of LSTM Network

3.3.2 Core Concept

The core concept of LSTM's are the cell state, and it's various gates. The cell state act as a transport highway that transfers relative information all the way down the sequence chain.The cell state, can carry relevant information throughout the processing of the sequence. So even information from the earlier time steps can make it's way to later time steps, reducing the effects of short-term memory. As the cell state goes on its journey, information get's added or removed to the cell state via gates. The gates are different neural networks that decide which information is allowed on the cell state. The gates can learn what information is relevant to keep or forget during training.

3.3.3 Gates

Gates control the flow of information to/from the memory. They controlled by a concatenation of the output from the previous time step and the current input and optionally the cell state vector.

3.3.3.1 Forget gate

First, LSTM have the forget gate. This gate decides what information should be thrown away or kept. Information from the previous hidden state and information from the current input is passed through the sigmoid function. Values come out between 0 and 1. The closer to 0 means to forget, and the closer to 1 means to keep.

3.3.3.2 Input Gate

To update the cell state, LSTM have the input gate. First, LSTM pass the previous hidden state and current input into a sigmoid function. That decides which values will be updated by transforming the values to be between 0 and 1. 0 means not important, and 1 means important. LSTM also pass the hidden state and current input into the tanh function to squish values between -1 and 1 to help regulate the network. Then LSTM multiply the tanh output with the sigmoid output. The sigmoid output will decide which information is important to keep from the tanh output.

3.3.3.3 Cell State

Now LSTM should have enough information to calculate the cell state. First, the cell state gets point wise multiplied by the forget vector. This has a possibility of dropping values in the cell state if it gets multiplied by values near 0. Then take the output from the input gate and do a pointwise addition which updates the cell state to new values that the neural network finds relevant. That gives new cell state.

3.3.3.4 Output Gate

Last gate is output gate. The output gate decides what the next hidden state should be. Remember that the hidden state contains information on previous inputs. The hidden state is also used for predictions.

First, pass the previous hidden state and the current input into a sigmoid function. Then pass the newly modified cell state to the tanh function. We multiply the tanh output with the sigmoid output to decide what information the hidden state should carry. The output is the hidden state. The new cell state and the new hidden is then carried over to the next time step

3.3.4 Tanh activation

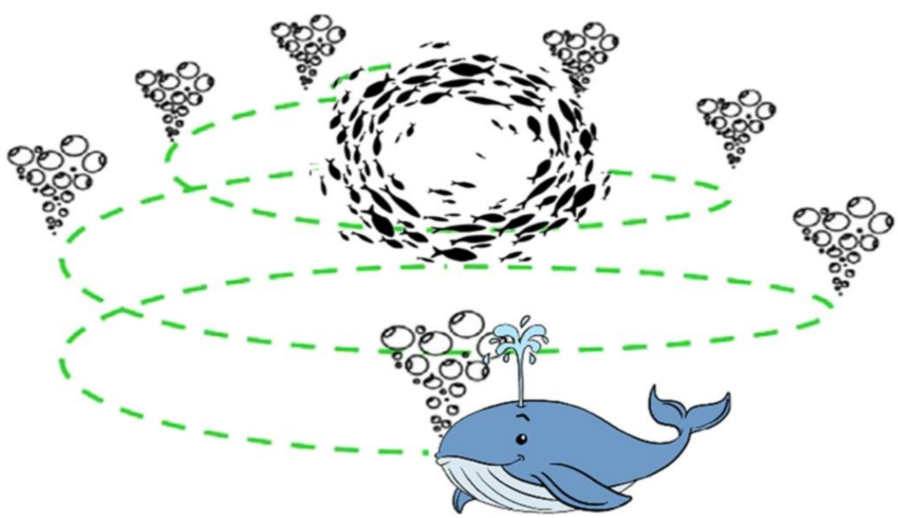
The tanh activation is used to help regulate the values flowing through the network. The tanh function squishes values to always be between -1 and 1.

3.3.5 Sigmoid

Gates contains sigmoid activations. A sigmoid activation is similar to the tanh activation. Instead of squishing values between -1 and 1, it squishes values between 0 and 1. That is helpful to update or forget data because any number getting multiplied by 0 is 0, causing values to disappears or be "forgotten." Any number multiplied by 1 is the same value therefore that value stay's the same or is "kept." The network can learn which data is not important therefore can be forgotten or which data is important to keep.

3.4 WHALE OPTIMIZATION ALGORITHM

Whale optimization algorithm was proposed by Jalili and Lewis for optimizing numerical problems (Mirjalili & Lewi, 2016). The algorithm simulates the intelligence hunting behavior of humpback whales. This foraging behavior is called bubble-net feeding method that is only be observed in humpback whales. The whales create the typical bubbles along a circle path while encircling prey during hunting. Simply, bubble-net hunting behavior could describe such that humpback whales dive down approximation 12 m and then create the bubble in a spiral shape around the prey and then swim upward the surface following the bubbles.



Whales are fancy creatures. They are considered as the biggest mammals in the world. An adult whale can grow up to 30 meters long and 180 tons weight. There are seven different main species of this giant mammal such as Killer, Minke, Sei, Humpback, Right, Finback, and Blue whales. Whales are mostly considered as predators. They never sleep because they have to breathe from the surface of oceans. In fact, half of the brain only sleeps. The interesting thing about the whales is that they are considered as highly intelligent animals with emotion.

According to Hof and Van Der Gucht [31], whales have common cells in certain areas of their brains similar to those of human called spindle cells. These cells are responsible for judgment, emotions, and social behaviors in humans. In other words the spindle cells make us distinct from other creatures. Whales have twice number of these cells than an adult human which is the main cause of their smartness. It has been proven that whale can think, learn, judge, communicate, and become even emotional as a human does, but obviously with a much lower level of smartness. It has been observed that whales (mostly killer whales) are able to develop their own dialect as well.

Another interesting point is the social behavior of whales. They live alone or in groups. However, they are mostly observed in groups. Some of their species (killer whales for instance) can live in a family over their entire life period. One of the biggest baleen whales is humpback whales (Megaptera novaeangliae). An adult humpback whale is almost as size of a school bus. Their favorite prey are krill and small fish herds.

The most interesting thing about the humpback whales is their special hunting method. This foraging behavior is called bubble-net feeding method [50]. Humpback whales prefer to hunt school of krill or small fishes close to the surface. It has been observed that this foraging is done by creating distinctive bubbles along a circle or '9'-shaped path. Before 2011, this behavior was only investigated based on the observation from surface.

However, Goldbogen et al. [20] investigated this behavior utilizing tag sensors. They captured 300 tag-derived bubble-net feeding events of 9 individual humpback whales. They found two maneuvers associated with bubble and named them 'upward-spirals' and 'double-loops'.

In the former maneuver, humpback whales dive around 12 meters down and then start to create bubble in a spiral shape around the prey and swim up toward the surface. The later maneuver includes three different stages: coral loop, lobtail, and capture loop. Detailed information about these behavior can be found in [20] .

It is worth mentioning here that bubble-net feeding is a unique behavior that can only be observed in humpback whales. In this work the spiral bubble-net feeding maneuver is mathematically modeled in order to perform optimization.In order to perform optimization, the mathematical model for spiral bubble-net feeding behavior is given as follows:

3.4.1 Encircling prey

Humpback whales can find the place of prey and encircle them. The WOA algorithm considers; current best search agent position be the target prey or close to the optimum point, and other search agents will try to update their position towards the best search agent. This behavior is formulated as the following equations:

$$D = |C.X^*(t) - X(t)| \qquad equ(3.1)$$

$$X(t+1) = X^*(t) - A.D \qquad equ(3.2)$$

where t indicates the current iteration, X∗ is the position vector of the best solution have been obtained so far iteration t, X is the position vector of each agent, | | is the absolute value, and . is an element-by-element multiplication. The coefficient vectors A and C are calculated as follows:

$$A = 2a.r - a \qquad equ(3.3)$$

$$C = 2r \qquad equ(3.4)$$

where a is linearly decreased from 2 to 0 over the course of the iteration and r is a random number [0,1].

3.4.2 Bubble-net attacking method

The Bubble-net strategy is hybrid of combined two approaches that can be mathematically model as follows:

3.4.2.1 Shrinking Encircling Mechanism

This behavior of Whales simulated by decreasing the value of a in the above equation. Note that the fluctuation range of A is also decreased by a . In other words, A is a random value in the interval [−a,a] where a is decreased from 2 to 0 over the course of iterations. Setting random values for A in [−1,1], the new position of a search agent can be defined anywhere in between the original position of the agent and the position of the current best agent.

3.4.2.2. Spiral Updating Position

In this approach, a spiral equation is created between the position of whale and prey to simulate the helix-shaped movement of humpback whales as follows:

$$D' = |X^*(t) - X(t)| \qquad equ(3.5)$$

$$X(t+1) = D'.e^{bl}.\cos(2\pi l) + X^*(t) \qquad equ(3.6)$$

where D′ is the distance between the whale and prey, b is constant defines the logarithmic shape, l is random in [−1,1] and is an element-by-element multiplication.

Indeed, humpback whales swim along a spiral-shaped path and at the same time within shrinking circle. Assuming a probability of 50%, choosing either the shrinking encircling movement or the spiral model movement is simulated during iterations of the algorithm. It means that:

$$X(t+1) = \begin{cases} X^*(t) - A.D, & if\ p < 0{,}5 \\ D'.e^{bl}.\cos(2\pi l) + X^*(t), & if\ p \geq 0.5 \end{cases} \qquad equ(3.7)$$

where p is a random number in [0,1].

3.4.3 Search for prey

Almost all meta-heuristic algorithms explore the optimum using random selection. In the bubble-net method, the position of the optimal design is not known, so humpback whales search for prey randomly. In contrast to the exploitation phase with A in interval [−1,1] in this phase consider, A be a vector of the random values greater than 1 or less than −1. With this assumption, search agent able to move far away from a reference whale. In return, the position of search agent will be updated according to randomly chosen from search agent, instead of the best search agent found so far. These two actions formulated as follows:

$$D = |C.X_{rand} - X| \qquad equ(3.8)$$

$$X(t+1) = X_{rand} - A.D \qquad equ(3.9)$$

where X_{rand} is a random position vector.

The WOA algorithm starts from a set of random solutions. At each iteration, search agents update their position according to the above explanations. WOA is a global optimizer. Adaptive variation of the search vector A, A allows the WOA algorithm easily transit between exploration and exploitation. Furthermore, WOA includes only two main internal parameters to be adjusted. High exploration ability of WOA is due to the position updating mechanism of Whales using (3.9). High exploitation and convergence are emphasized, which originate from (3.6) and (3.2). These equations show that the WOA algorithm is able to provide high local optima avoidance and convergence speed during the course of the iteration.

4. RESULTS AND DISCUSSION

This section presents the simulation results of the proposed method for loan eligibility prediction. Then, the performance is analyzed with such metrics are precision,accuracy and mean-square-error.

4.1 Dataset description

The dataset contains Dream Housing Finance company deals in all home loans. They have presence across all urban, semi urban and rural areas. This data set contain the details about the loan candidate. The details are, Gender, Marital Status, Education, Number of Dependents, Income, Loan Amount, Credit History and others.

4.2 Experimental set up

The proposed method of loan eligibility prediction is implemented using PYTHON under Pycharm framework with tensorflow backend.

4.3 Evaluation metrics

The performance of the proposed method is analyzed by accuracy, mse and precision. These metrics are explained as follows:

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

$$Precison = \frac{TP}{TP + FP}$$

$$MSE = \frac{1}{n} \sum_{n=1}^{n} (original\ class_i - predicted\ class_i)$$

Where, TN denotes True Negative, TP indicates True Positive, FN denotes False Negative and FP signifies False Positive.

4.4 Comparative analysis

Figure 3 shows the performance analysis of credit card fraud detection methods based on various percentage of training data. While using the percentage of training data is 50, the existing neural network achieved 60% precision. However, the proposed deep neural network acquires the higher value of 72.5%. Thus, the proposed algorithm attains higher precision value of 90% when compared with the neural network in 70% of training data. Similarly, the performance analysis of accuracy is shown in figure 4. The existing neural network obtains 67.43% while using the percentage of training data is 50 but the proposed method achieves the accuracy of 70%. Rather than the existing method, the proposed algorithm achieves 90% of accuracy value for 70% of training data.

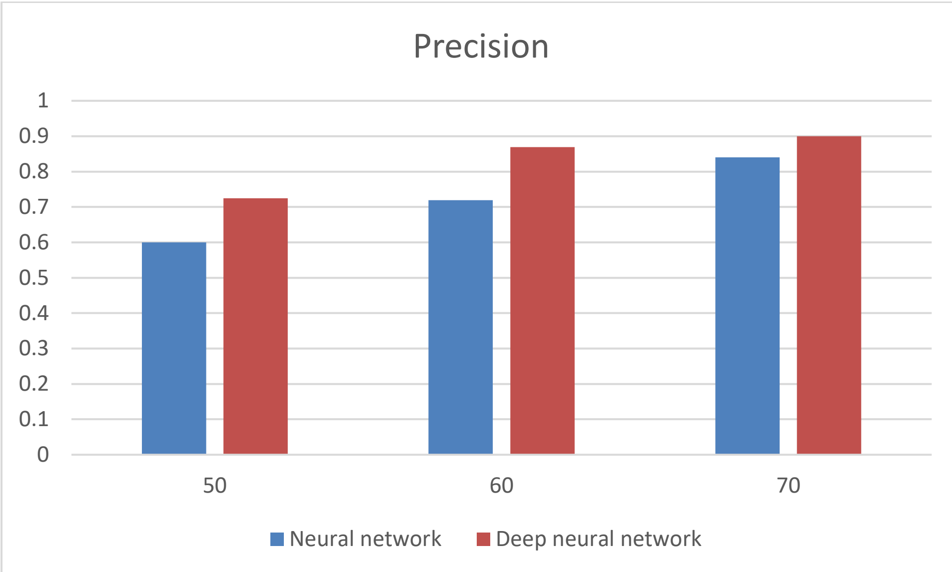


Fig.3. Comparative analysis based on precision

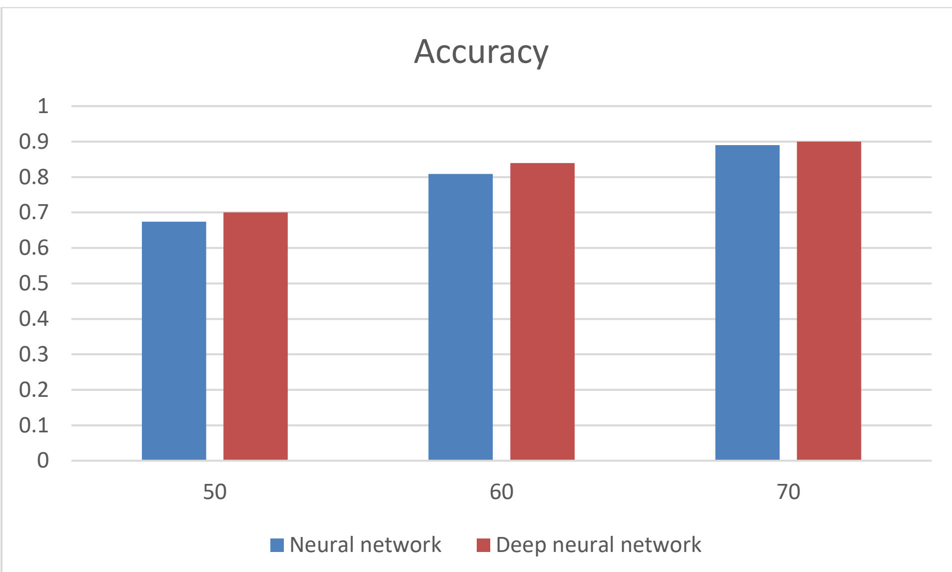


Fig.3. Comparative analysis based on accuracy

5. CONCLUSION

The alarming increase in fraudulent credit card usage has stressed the fraud management systems currently in use in banks and other institution that process credit card transaction. To progress safety measures of the monetary transaction systems in a habitual and effectual way, structure a precise and

(IJAER) 2021, Vol. No. 22, Issue No. I, July **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

well organized credit card scam detection system is one of the essential functions for money transactions. Accordingly, deep neural network based method is developed in this paper for credit card fraud detection. Here, four important steps such as, 1) Payment Request terminal, 2) Request based feature extraction, 3) Feature augmentation, 4) Deep Neural Network is used for the detection of fraudulent behavior of the credit card users. Finally, the experimentation is performed with benchmark dataset and the results proved that the proposed method attained the accuracy of 90% which is higher when compared with the existing neural network. In future, the optimization methods can be applied for improvement of training process of deep neural network.

REFERENCES

1. Awoyemi, J. O., Adetunmbi, A. O. and Oluwadare, S. A. (2017) 'Credit card fraud detection using machine learning techniques: A comparative analysis', Proceedings of the IEEE International Conference on Computing, Networking and Informatics, ICCNI 2017, 2017–January, pp. 1–9. doi: 10.1109/ICCNI.2017.8123782.
2. Barman, D., Chowdhury, N. and Singha, R. K. (2012) 'To predict possible profit/loss of a movie to be launched using MLP with back-propagation learning', Proceedings of the 2012 International Conference on Communications, Devices and Intelligent Systems, CODIS 2012, pp. 322–325. doi: 10.1109/CODIS.2012.6422203.
3. Benchaji, I., Douzi, S. and ElOuahidi, B. (2019) 'Using Genetic Algorithm to Improve Classification of Imbalanced Datasets for Credit Card Fraud Detection', pp. 1–5. doi: 10.1109/csnet.2018.8602972.
4. Bendangnuksung and Prabu, D. (2018) 'Students' Performance Prediction Using Deep Neural Network', International Journal of Applied Engineering Research, 13(2), pp. 1171–1176. Available at: http://www.ripublication.com.
5. Carneiro, E. M. et al. (2015) 'Cluster Analysis and Artificial Neural Networks: A Case Study in Credit Card Fraud Detection', Proceedings - 12th International Conference on Information Technology: New Generations, ITNG 2015, pp. 122–126. doi: 10.1109/ITNG.2015.25.
6. Chandrasekar, P., & Qian, K. (2016). The Impact of Data Preprocessing on the Performance of a Naive Bayes Classifier. 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC). doi:10.1109/compsac.2016.205.
7. Dey, A. (2016) 'Machine Learning Algorithms: A Review', International Journal of Computer Science and Information Technologies, 7(3), pp. 1174–1179. Available at: www.ijcsit.com.
8. European Union (2009) 'European Union Risk Assessment Report:TRIS (2-CHLOROETHYL) PHOSPHATE, TCEP', (July), p. 213.
9. Ghosh and Reilly (1994) 'NN02: Credit card fraud detection with a neural-network', Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences HICSS-94, pp. 621–630. doi: 10.1109/HICSS.1994.323314.

10. Halim, S. A. et al. (2011) 'A comparative study between standard Back Propagation and Resilient Propagation on snake identification accuracy', ITME 2011 - Proceedings: 2011 IEEE International Symposium on IT in Medicine and Education, 2, pp. 242–246. doi: 10.1109/ITiME.2011.6132031.
11. Hormozi, H. et al. (2013) 'Credit cards fraud detection by negative selection algorithm on hadoop', IKT 2013 - 2013 5th Conference on Information and Knowledge Technology, pp. 40–43. doi: 10.1109/IKT.2013.6620035.
12. http://www.efxkits.co.uk (1973) 'New Text Document', Phil. Mag., p. 1193. doi: 10.1080/15228951003772454.
13. Hua, J. (2008). Study on the Application of Rough Sets Theory in Machine Learning. 2008 Second International Symposium on Intelligent Information Technology Application. doi:10.1109/iita.2008.154
14. Iyer, D. et al. (2011) 'Credit card fraud detection using hidden Markov model', Proceedings of the 2011 World Congress on Information and Communication Technologies, WICT 2011, 5(1), pp. 1062–1066. doi: 10.1109/WICT.2011.6141395.
15. Kaur, S. and Deol, R. (2017) 'Naive Bayes Classification Technique for Opinion Mining in Data Mining', 2(3), pp. 44–47.
16. Kesavaraj, G. and Sukumaran, S. (2013) '06726842'. doi: 10.1109/ICCCNT.2013.6726842.
17. Lan, K. et al. (2018) 'A Survey of Data Mining and Deep Learning in Bioinformatics', Journal of Medical Systems, 42(8). doi: 10.1007/s10916-018-1003-9.
18. Lei, L. (2012) 'Card fraud detection by inductive learning and evolutionary algorithm', Proceedings - 2012 6th International Conference on Genetic and Evolutionary Computing, ICGEC 2012, (156), pp. 384–388. doi: 10.1109/ICGEC.2012.70.
19. Li, Z. et al. (2019) 'Credit Card Fraud Detection Using Capsule Network', 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, pp. 3679–3684. doi: 10.1109/smc.2018.00622.
20. Ma, W. et al. (2018) 'Construction of Layered Structure of Anion-Cations to Tune the Work Function of Aluminum-Doped Zinc Oxide for Inverted Polymer Solar Cells', ACS Applied Materials and Interfaces, 10(12), pp. 10513–10519. doi: 10.1021/acsami.7b16653.
21. Malini, N. and Pushpa, M. (2017) 'Analysis on credit card fraud identification techniques based on KNN and outlier detection', Proceedings of the 3rd IEEE International Conference on Advances in Electrical and Electronics, Information, Communication and Bio-Informatics, AEEICB 2017, pp. 255–258. doi: 10.1109/AEEICB.2017.7972424.
22. O'Shea, K. and Nash, R. (2015) 'An Introduction to Convolutional Neural Networks', (November). doi: 10.1007/978-3-642-28661-2-5.
23. Pannu, A. and Student, M. T. (2008) 'Artificial Intelligence and its Application in Different Areas', international Journal of Engineering and Innovative Technology, 4(10), pp. 79–84. doi: 10.1155/2009/251652.

24. Patel, B. N. (2012) ‘Efficient Classification of Data Using Decision Tree’, Bonfring International Journal of Data Mining, 2(1), pp. 06-12. doi: 10.9756/bijdm.1098.
25. Paul, L. C., Suman, A. Al and Sultan, N. (2013) ‘Methodological Analysis of Principal Component Analysis (PCA) Method’, IJCEM International Journal of Computational Engineering & Management ISSN, 16(2), pp. 2230–7893.
26. Pozzolo, A. D. et al. (2015) ‘Credit Card Fraud Detection and Concept-Drift Adaptation with Delayed Supervised Information’.
27. Sarkar, J., Saha, S. and Agrawal, S. (2014) ‘An Efficient Use of Principal Component Analysis in Workload Characterization-A Study’, AASRI Procedia. Elsevier B.V., 8(Secs), pp. 68–74. doi: 10.1016/j.aasri.2014.08.012.
28. Sehgal, S. et al. (2014) ‘Data analysis using principal component analysis’, 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems, MedCom 2014, (2), pp. 45–48. doi: 10.1109/MedCom.2014.7005973.
29. Sheikh, R. H., Raghuwanshi, M. M. and Jaiswal, A. N. (2008) ‘Genetic algorithm based clustering: A survey’, Proceedings - 1st International Conference on Emerging Trends in Engineering and Technology, ICETET 2008, 2(6), pp. 314–319. doi: 10.1109/ICETET.2008.48.
30. ShiZhen and ZhouYang, C. T. (2005) ‘Comparison of steady state and elitist selection genetic algorithms’, (August), pp. 495–499. doi: 10.1109/icima.2004.1384245.
31. Soltani, N., Akbari, M. K. and Sargolzaei Javan, M. (2012) ‘A new user-based model for credit card fraud detection based on artificial immune system’, AISP 2012 - 16th CSI International Symposium on Artificial Intelligence and Signal Processing, (Aisp), pp. 29–33. doi: 10.1109/AISP.2012.6313712.
32. Srivastava, A. et al. (2016) ‘Credit Card Fraud Detection at Merchant Side using Neural Networks’, Proceedings of the 10Th Indiacom - 2016 3Rd International Conference on Computing for Sustainable Global Development, pp. 667–670.
33. Stephen Marsland, “Machine Learning – An Algorithmic Perspective”, Second Edition, Chapman and Hall/CRC Machine Learning and Pattern Recognition Series, 2014.
34. Thengade, A. (2012) ‘Genetic Algorithm: Survey Paper’, International Journal of Computer Applications, 2012(January 2012), pp. 7–8.
35. Thomas, T., Bhuvan, N. T. and Professor, A. (2017) ‘Study on Distributed SDN Controllers and Failover Mechanisms’, International Journal of Innovative Research in Computer and Communication Engineering, 5(4), pp. 11614–11619. doi: 10.15680/IJIRCCE.2017.
36. Tongur, V. and Ülker, E. (2014) ‘Migrating Birds Optimization for Flow Shop Sequencing Problem’, Journal of Computer and Communications, 02(04), pp. 142–147. doi: 10.4236/jcc.2014.24019.
37. Tu, P.-L. and Chung, J.-Y. (2003) ‘A new decision-tree classification algorithm for machine learning’, pp. 370–377. doi: 10.1109/tai.1992.246431.

38. Wang, C. et al. (2018) 'Credit card fraud detection based on whale algorithm optimized BP neural network', 13th International Conference on Computer Science and Education, ICCSE 2018. IEEE, (Iccse), pp. 614–617. doi: 10.1109/ICCSE.2018.8468855.
39. Chen, Qiao-ling, and Jian-bang Lin," Integrating of business intelligence and CRM in banks: An empirical study of SOM applied in personal customer loans in Taiwan," Fuzzy Theory and Its Applications (iFUZZY), 2015 International Conference on. IEEE, 2015
40. Y. Dong, M.Firth,W. Hou, W. Yang , " Evaluating the performance of Chinese commercial banks: A comparative analysis of different types of banks,"European Journal of Operational Research, pp.280-295, 2016
41. E.Ozsoz., E.W Rengifo,M.A Akinkunmi, "What determines return risks for bank equities in Turkey?," Borsa Istanbul Review ,pp.23-31,2014
42. S.Mammadli,"Fuzzy Logic Based Loan Evaluation System," Procedia Computer Science, , pp.495-499, 31 Dec 2016
43. C. Xiaojie, D.Huailin, W.Qingfeng,"SCAD algorithm and its application in analyse of bank loan," Computer Science & Education, 2009. ICCSE'09. 4th International Conference, pp. 1934-1939, July 2009
44. V.Kumar,Natarajan.S,Keerthana,Chinmayi K.M. and Lakshmi Lupu ,"Credit Risk Analysis in Peer-to-Peer Lending System, " Knowledge Engineering and Applications (ICKEA), IEEE International Conference on IEEE , pp. 193-196, September 2012
45. Xudong Lin, Lin Cheng, W.Mao, Z.Qiu, "Research on measuring the bank's loan willingness based on prospect theory," Service Systems and Service Management (ICSSSM), 12th International Conference on IEEE ,p.p 1-5,June 2015
46. A.Calis, A.Boyaci, K.Baynal,"Data mining application in banking sector with clustering and classification methods, " Industrial Engineering and Operations Management (IEOM),International Conference on IEEE , p.p 1-8,March 2015
47. Yu Zhang, Guang Yu, Yong-sheng Guan, and Dong-hui Yang, "Feature selection of nonperforming loans in Chinese commercial banks,"Management Science & Engineering (ICMSE), 2014 International Conference on IEEE , pp. 1208-1215, 2014
48. Zenglian Zhang,"Research of default risk of commercial bank's personal loan based on rough sets and neural network," Intelligent Systems and Applications (ISA), 3rd International Workshop on IEEE,p.p 1-4, 2011
49. Liu Bingwu,Li Zhou e"The Study of Commercial Bank Loan Risk Pricing Based on RAROC Model," International Information Management, Innovation Management and Industrial Engineering, vol.2, 2009
50. Shiivong Birla, Kashish Kohli, Akash Dutta ,"Machine Learning on imbalanced data in Credit Risk," Information Technology, Electronics and Mobile Communication Conference (IEMCON), p.p 1-6, 2016

51. Chao Wu, and Wenhe Li,"Analysis of Interbanking Systemic Risk Contagion-Case from Finland Bank System,"Social and Management of e-Commerce and e-Government (ICMeCG), p.p 82-86, 2014
52. Yang Yanbing, Liu Bin ," Bayes' Rule and Bank Credit Game Risk Analysis," Measuring Technology and Mechatronics Automation (ICMTMA) , pp. 335-337,2013
53. Yinxiao Ma,Hong LiuS.Vos, J.Scheerder ," Research of SVM Applying in the Risk of Bank's Loan to Enterprises," Information Engineering and Computer Science (ICIECS) ,p.p 1-5,2015
54. P. R. Hof and E. Van Der Gucht, "Structure of the cerebral cortex of the humpback whale, Megaptera novaeangliae (Cetacea, Mysticeti, Balaenopteridae)," The Anatomical Record, vol. 290, pp. 1-31, 2007.
55. W. A. Watkins and W. E. Schevill, "Aerial observation of feeding behavior in four baleen whales: Eubalaena glacialis, Balaenoptera borealis, Megaptera novaeangliae, and Balaenoptera physalus," Journal of Mammalogy, pp. 155-163, 1979.
56. J. A. Goldbogen, A. S. Friedlaender, J. Calambokidis, M. F. Mckenna, M. Simon, and D. P. Nowacek, "Integrative Approaches to the Study of Baleen Whale Diving Behavior, Feeding Performance, and Foraging Ecology," BioScience, vol. 63, pp. 90-100, 2013.
57. M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep lstm-rnn," Neural Computing and Applications, pp. 1–14, 2017.