



ABC_LSTM: Optimizing Parameters of Deep LSTM using ABC Algorithm for Big Datasets

Shweta Mittal, Om Prakash Sangwan

Abstract: Long Short Term Memory Network is the variant of RNN (Recurrent Neural Network) popularly used in various domains, particularly for sequence prediction tasks. For deep networks, number of hidden layers in the network is high and thus, the time complexity of the network increases. Moreover, with the increase in the size of datasets, it becomes very difficult to tune these complex networks manually (as the network may take several days/weeks to run). Thus, to minimize the time required to run an algorithm and for better accuracy, there is a need to automate the task of tuning the parameters of the network. To automatically tune the parameters of the networks, various researchers have used numerous Metaheuristic approaches like Ant Colony Optimization, Genetic Algorithm, Simulated Annealing etc. in the past which provides us with the near optimal solution. In the proposed ABC_LSTM algorithm, traditional Artificial Bee Colony algorithm has been implemented to optimize the number of hidden neurons of LSTM networks with 2 hidden layers. Based on the experimental results, it can be concluded that up to a certain point increasing the number of bees and iterations gives us the solution with the least MAE value, thereby improving the accuracy of the model.

Keywords : Deep Learning, LSTM, Big Data, ABC.

I. INTRODUCTION

Big data is a term defined for huge amount of data generated at a very high speed containing values of different data types revealing some important information. Telecom industry, Banking sector, Social media data, Weblogs, Meteorological data are the various sources of Big data. Various researchers are exploring tools like Spark, Storm, Cassandra, Hadoop, YARN, MongoDB, GPUs etc. which can effectively store and process big datasets. Among these above mentioned tools, Spark is a very popular open source tool developed by Apache for big datasets, which has an advantage of in-memory data processing and support for iterative tasks.

Spark has basically five components: i) Spark Core responsible for memory management, task scheduling, fault

recovery etc. ii) Spark Streaming: a component for data stream processing iii) Spark SQL: a package for querying data via SQL, Hive etc. iv) GraphX library for graph related computations and v) MLlib library to support numerous machine learning algorithms. Any spark application can be run in following modes: either locally, or in cluster mode on Stand-alone, YARN or Mesos platform (which schedules the resources among various nodes in the cluster). Whenever an application is run, Spark Driver uses Spark Context to connect to the cluster while Spark Session provides an interface to that connection. Creating more than one instance of SparkSession at a time might create an issue. Therefore, to avoid the problem, SparkSession.builder.getOrCreate() method should be used which returns an instance if it exists already, else creates a new one. Spark Configuration allows a user to specify number of cores per executor (spark.executor.cores), number of executor instances (spark.executor.instances), amount of memory allocated to driver (spark.driver.memory) or executor (spark.executor.memory) etc.

RDD i.e. Resilient Distributed Datasets is the basic data structure in Spark which is immutable, fault tolerant and supports lazy operations. Whenever a file is loaded, RDD splits the file into multiple partitions across the cluster which is then executed in parallel. The number of partitions have default value i.e. spark.default.parallelism where default is the number of cores in the cluster. Dataframe is the high abstraction data structure build on the top of RDDs which are highly optimized by Project Tungsten Memory Management and Catalyst Optimizer. Pandas is another open source library which provides dataframes that supports in-memory computing and is limited to the size of server.

LSTM RNN i.e. Long Short Term Memory Recurrent Neural Network is the most popular architecture of neural network which is composed of a memory unit and a feedback loop. LSTM network has following components: a cell, an input gate, an output gate and a forget gate. LSTM is widely used for sequence prediction task in number of domains, particularly in tasks to evade the dependency concern. Deep Learning is an area in the field of machine learning algorithms which has multiple hidden layers to obtain features from raw data. In this proposed work, LSTM network with 2 hidden layers has been implemented.

Metaheuristic techniques are nature inspired search techniques which provide us with the near optimal solution and can be classified as described in Fig. 1 [1]. Artificial Bee Colony, Particle Swarm, Ant Colony,

Revised Manuscript Received on May 15, 2020.

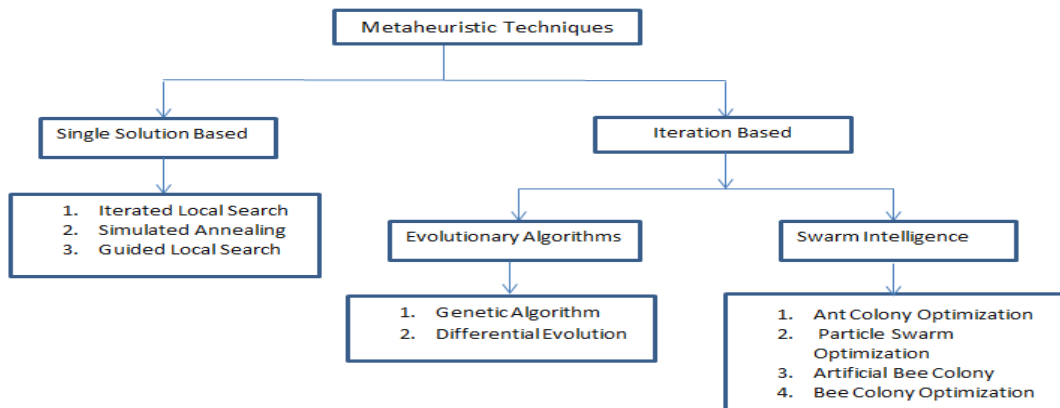
* Correspondence Author

Shweta Mittal*, Deptt. Of Computer Science, Guru Jambheshwar University of Science and Technology, Hisar, Haryana. Email: shwetamittal019@gmail.com

Om Prakash Sangwan, Deptt. Of Computer Science, Guru Jambheshwar University of Science and Technology, Hisar, Haryana. Email: sangwan0863@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Genetic Algorithm are the various Metaheuristic techniques which are gaining immense popularity these days. Among these techniques, Artificial Bee Colony has been widely used in numerous domains as it has the capability of finding global optimum solution.



ABC algorithm was presented by Dervis Karaboga in 2005 to minimize the objective function and is motivated from the foraging characteristics of the bees. It comprises of three types of bees: Employed, Onlooker and Scout bees. Employed bees search for food sources in the environment which then communicate the information collected to the onlooker bees in the form of waggle dance. Thereafter, Onlooker bees choose the food source with some probability selection method (i.e. by Greedy selection/ Roullete wheel selection method). The bee which has the highest happiness value/fitness value has the maximum chance of getting selected. Scout bees are the bees which are not selected after some predefined number of trials and are abandoned from the search process.

II. LITERATURE REVIEW

In our previous work, we have reviewed various machine learning and data mining techniques used for analysis of Big Data and it was interpreted that Deep Learning, Online Learning and Incremental Learning are popular techniques used in the field of big data[2], [3]. Also, we have successfully implemented LSTM RNN algorithm to analyze 15 years of meteorological big data on Spark platform [4]. In the past few decades, various researchers have automated the task of tuning the parameters of the Neural Networks and the same has been reviewed in this section.

H. Nimbark *et al.* [5] optimized architecture of RNN using improved version of ABC algorithm with the help of chaos based initialization and the proposed algorithm proved to be robust and flexible. B.A. Garro *et al.* [6] presented a methodology to optimize the parameters of ANN such as initial weights, transfer functions and number of neurons in the hidden layer using different versions of PSO and compared the performance with LM and Backpropagation algorithm. X. Cai *et al.* [7] implemented RNN using hybrid PSO-EA algorithm and the proposed algorithm gave the satisfactory results. A.K. Tripathy *et al.* [8] optimized ANN using PSO for weather forecasting task using three years of weather data.

N.M. Nawi *et al.* [9] implemented Cuckoo search for

weight optimization of RNN and it was concluded that CSERN (Cuckoo Search Elman Recurrent Network) and CSBPERN (Cuckoo Search Back Propagation Elman Recurrent Network) performs superior as compared to ABC-BP, ABC-LM, ABCNN and BPNN algorithm in terms of accuracy and MAE. H.A. Albeahdili *et al.* [10] optimized parameters of CNN via PSO algorithm hybrid with GA on MNIST, CIFAR-10, and SVHN dataset. To avoid the shortcoming of PSO of getting stuck in local optima, GA was used. A hybrid training algorithm was presented by Q. Kang *et al.* [11] using RNN, PSO and Temporal Error Aggregation in which changes were made in the computation of objective error in the backpropagation stage.

Tiantian Li *et al.* [12] utilized 24 hours of load data to forecast the load for the next day using EMD-PSO-LSTM and the algorithm proved to be quite effective. I. Salman *et al.* [13] optimized hyper-parameters of LSTM using DE, PSO, Simulated Annealing and Random Search algorithm for emotion recognition application.

Salman *et al.* [14] studied impact of metaheuristic iterations on ANN and the results concluded that GA overcome PSO and Fireworks algorithm. H. Chiroma *et al.* [15] surveyed progress of ANN for analysis of big data and it was found that Deep Convolution Net, Parallel BPNN, De-centralised RNN, Swarm Algorithms have been used by various researchers for analysis of Big Data. In order to increase the efficiency of BPNN, T.A. Rashid *et al.* [16] combined Artificial Bee algorithm with the crossover operation of Genetic algorithm.

Instead of optimizing the parameters, O. Zavalnyi *et al.* [17] optimized the gradients using Metaheuristic techniques while training of Deep NN. Experiment was conducted on MNIST and CIFAR-10 dataset with population size of 5, 50 while SGD, Adam and RMSProp optimizers have been used for evaluation purpose. From the results, it was concluded that Adam and RMSProp adapts favorable gradients more efficiently than SGD.

J. Zhang *et al.* [18] implemented GADAM i.e. Genetic Evolutionary ADAM which learns unit model with Adam and evolved using GA. Parallelization is achieved using separate processors/servers and it was analyzed that with increase in number of generations, performance is first improved and then becomes stable. Further, if population size is large, average learning performance may vary. A.P. Pitowski *et al.* [19] compared various Metaheuristic techniques for Artificial Neural Network training to predict water temperature and it was concluded that LM algorithm outperforms various Metaheuristic approaches.

E. Nourani *et al.* [20] proposed hybrid ABC based Neural Network in which roulette wheel selection is used for onlooker bees instead of greedy selection and the results concluded that proposed hybrid approach performs promising in comparison to Genetic Algorithm, Artificial Bee Colony, Gradient Decent and LM algorithm. J.M. Ordonez *et al.* [21] implemented a framework for training of deep NNs using Tensorflow framework on both CPU and GPU which also allows easy integration of various Metaheuristic techniques.

X.Y. Chen *et al.* [22] performed comparative study of various Metaheuristic algorithm (i.e. DE, ABC and ACO algorithm while PSO is used as standard for evaluation) for training of Artificial Neural Network and it was interpreted that Differential Evolution and Ant Colony algorithm are more compatible while ABC has slow convergence and is very time consuming.

III. PROPOSED ABC_LSTM ALGORITHM

From the review done in the above section, it is concluded that various Metaheuristic algorithms like PSO, GA, ABC, ACO etc. have been successfully used to optimize the parameters of various Neural Networks. In this paper, we have proposed hybrid ABC_LSTM algorithm in which LSTM RNN neural network is used for prediction of temperature for 1 hour time lag while ABC algorithm is used to automate the choice of number of hidden layers for the LSTM network. We chose to hybrid both these techniques as LSTM is good for exploitation only and can get trapped in local optima while ABC algorithm is good for exploration, thus hybrid of these algorithm helps in achieving good balance between exploration and exploitation. The steps for the proposed hybrid algorithm have been defined below:

A. Initialize the number of food sources say 2, 5 or 10. The length of each food source is determined by the number of parameters we wish to optimize. The number of hidden layers for the proposed network has been set to 2, thus there are 2 parameters i.e. number of neurons in the 2 hidden layers that needs to be optimized. Upper and lower bound of both the parameters has been set to 1 and 50 respectively. The higher will be the bound, more will be the neurons in hidden layer and larger will be the complexity of the network. Employed bees in a hive are equal to strength of onlooker bees, and also to the number of food sources. For 5 food sources, initial population will be generated as follows:

[(5,15), (39,1), (45,25), (12,35), (1,27)]

Thus, Employed bees=Onlooker bees=5. Here, (5,15) represents 5 neurons in the first hidden layer and 15 neurons in the second hidden layer respectively.

B. Repeat steps C. to E. for n number of iterations

C. For each employed bee, a new food source is generated by crossover and mutation operator. If new food source's fitness/happiness value is larger than previous one, then the new solution vector is allocated to the corresponding employed bee. Happiness value/Fitness value for every food source is calculated via below equation:

$$\text{Fitness_value}[i] = 1/(1+\text{value}[i]) \quad (1)$$

where value[i] of food source is MAE value evaluated by the Deep LSTM RNN network.

Employed bees then performs the waggle dance and communicates the information collected to the onlooker bees.

D. Each onlooker bee will then try to improve the food source locally by some probability selection method, which is roulette wheel selection in our case. Roulette wheel selection method is used to find out the best food source among previous food sources and the one selected by onlooker bee. Probability of a given food source to be selected by the onlooker bee is determined by the following equation. :

$$\text{Prob_value}[i] = [0.9 * \text{fitness}[i] / \text{Total_fitness}] + 0.1 \quad (2)$$

where Total_fitness is the sum of fitness of all the food sources in the hive.

E. If bees are not getting better after some predefined number of trials, then it is replaced by some new randomly generated food source and the corresponding bees are termed as scout bees.

IV. EXPERIMENTAL RESULTS

For the proposed hybrid algorithm, 15 years of meteorological data collected from across 122 weather stations for south-eastern states of Brazil has been given as input to LSTM RNN Neural network to predict temperature for 1 hour time lag (parameters of which has been optimized using ABC algorithm). Dataset is around 2 GB and can be downloaded from Kaggle.com [23]. The whole experiment is implemented on a computer with 1 TB memory and 16 GB RAM. Proposed ABC_LSTM algorithm is implemented in python library on Spark platform. Details for implementation of LSTM algorithm can be found in paper [4].

Number of food sources/employed bees has been set to 5, 10 and 20 while number of iterations has been set to 2, 5 and 10 respectively. Table 1 highlights the best choice of number of hidden neurons obtained by the proposed ABC_LSTM algorithm for 8 different runs of algorithm, their corresponding MAE value (i.e. Mean Absolute Error) and the time taken by each run. The best result has been achieved for network having 3 hidden neurons in first layer and 8 hidden neurons in second layer respectively and the algorithm took around 130 hours i.e. 5.4 days to run.

From Table 1, it can also be inferred that for 5 employed bees, MAE value of proposed algorithm decreases from 2 to 5 iterations (i.e. from 0.0161 to 0.015966) and then increases slightly for 10 iterations (i.e. from 0.015966 to 0.015969, as shown in Fig. 2). While for 10 employed bees, MAE value has a continuous decreasing curve (i.e. from 0.016001 to

0.015876 and to 0.015713, as shown in Fig. 3). Also, with increase in the number of employed bees, MAE value of the network decreases i.e. from 0.161 to 0.016001 and to 0.015963 (when algorithm is run till 2 iterations).

Table- I : MAE value and time taken for 5, 10 and 20 employed bees for 8 different runs

No of bees	No of iterations	Best Soln	MAE	Time(sec)	Time(hours)	Time(days)	No. of times eval is run
5	2	39,1	0.0161	15255.3	4.237583333	0.176565972	30
5	5	9,1	0.015966	75275.88019	20.90996672	0.871248613	68
5	10	1,32	0.015969	485010.4564	134.7251268	5.613546949	134
10	2	32,6	0.016001	21915.46257	6.087628493	0.253651187	50
10	5	3,2	0.015876	361101.7991	100.3060553	4.179418971	107
10	10	3,8	0.015713	469307.0576	130.3630715	5.431794648	214
20	2	2,1	0.015963	92293.51244	25.63708679	1.06821195	102
20	3	9,2	0.015948	183474.1714	50.96504762	2.123543651	142

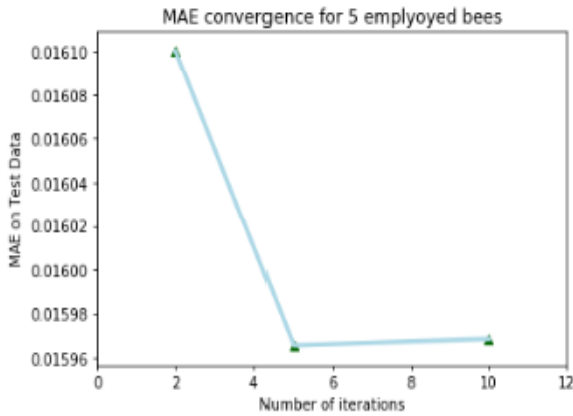


Fig. 2. MAE convergence for 5 employed bees runs for 2,5 and 10 iterations

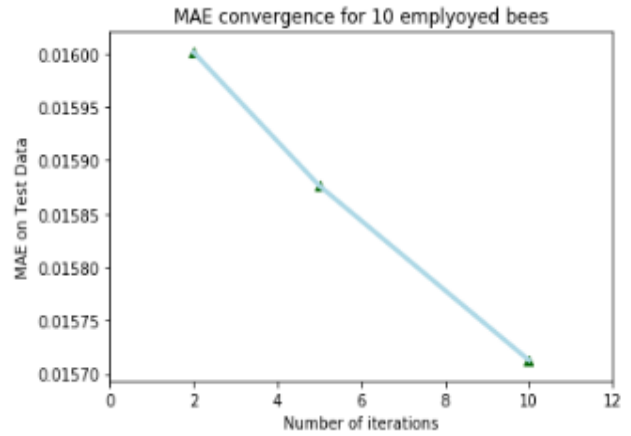


Fig. 3. MAE convergence for 10 employed bees for 3 different runs for 2, 5 and 10 iterations

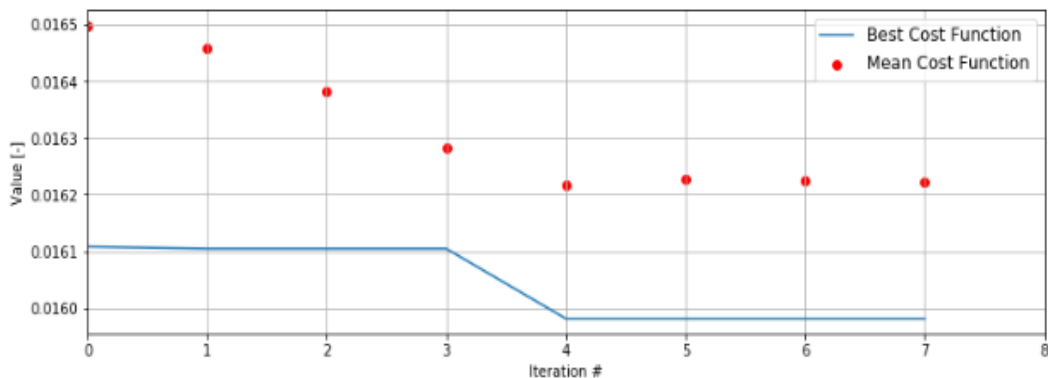


Fig. 4. Best cost and mean cost value for 5 employed bees from iterations 1 to 8 for a single run

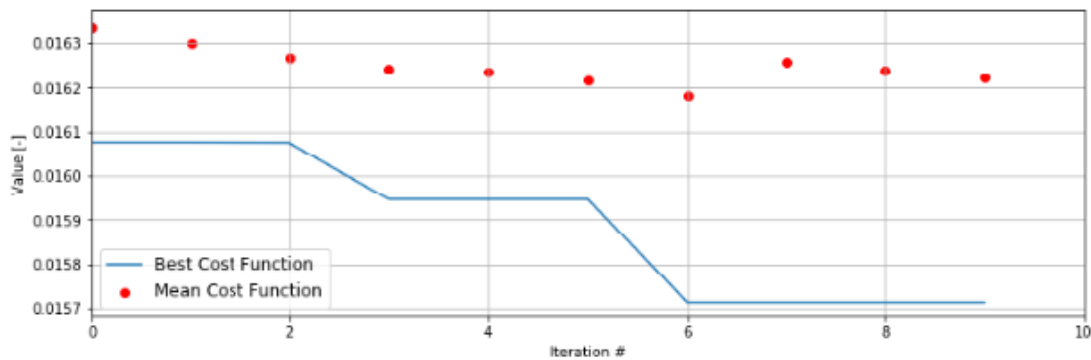


Fig. 5. Best cost and mean cost value for 10 employed bees from iterations 1 to 10 for a single run

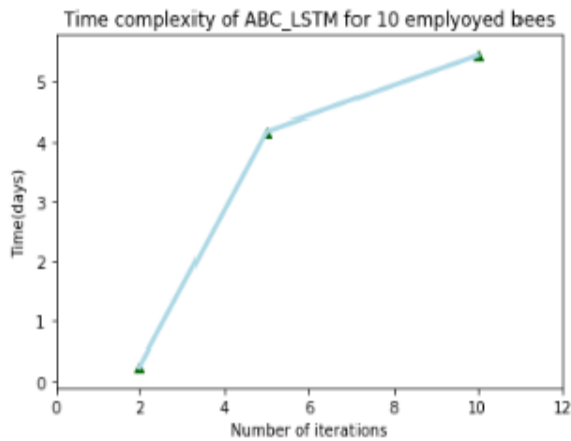


Fig. 6. Time complexity of ABC_LSTM for 10 employed bees over iterations

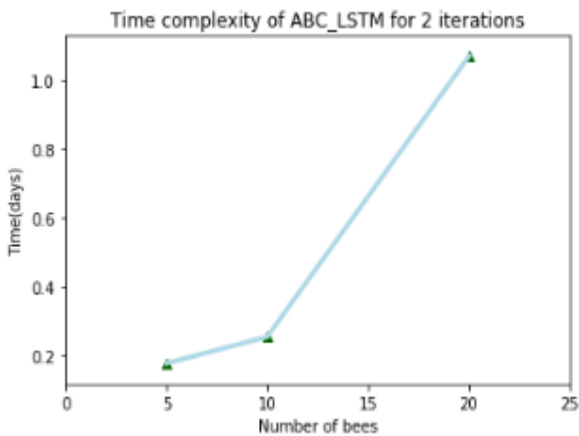


Fig. 7. Time complexity of ABC_LSTM for 2 iterations over number of bees

Fig. 4 and 5 represents the best cost value (i.e. least MAE value of all the bees) and mean cost value (i.e. mean of the MAE value of all the bees) for the proposed algorithm over generations/iterations for a single run. Here, red dot represents the mean cost value while blue curve represents the best cost value for that particular iteration. On x axis, 0 represents 1st iteration, 1 represents 2nd iteration and so on while y axis represents the MAE value. Fig. 4 represents the case study of ABC_LSTM algorithm having 5 employed bees run until 8 iterations and it can be interpreted that after 5 iterations, both best cost and mean value change only slightly and thus we can say that algorithm converges after 5

iterations. Similarly, Fig. 5 represents the case study having 10 employed bees run until 10 iterations and it can be interpreted that algorithm converges after 7 iterations.

Fig. 6 represents the time complexity of the proposed algorithm for 10 employed bees run till 2,5 and 10 iterations (for 3 different runs of proposed algorithm) and it is inferred that time complexity for the proposed algorithm increases marginally with increase in iterations. As number of iterations increases, the number of times evaluation function is run also increases and thus the time complexity increases. Similarly, Fig. 7 represents the time complexity of proposed algorithm run till 2 iterations for 2, 5 and 10 employed bees (for 3 different runs) and it can be concluded that time complexity also increases marginally with increase in number of employed bees.

V. CONCLUSION AND FUTURE WORK

In this paper, we have proposed ABC_LSTM algorithm to optimize the number of hidden neurons for the Deep LSTM network using traditional ABC algorithm as it is difficult to tune the parameters of network manually. Time complexity of our proposed model with increase in the number of bees and number of iterations has also been analyzed. From the experimental study, it is interpreted that with increase in number of iterations of the proposed algorithm, MAE value of the network first decreases up to the certain number of iterations and then becomes constant. Thus, increasing the number of iterations after a certain point, there will be no further improvement in the MAE value and hence, the algorithm converges. Also, we can observe that MAE value of the LSTM network decreases with increase in the number of bees. In this experimental study, the upper limit for the number of employed bees has been set to 20 and maximum iteration value has been set to 10 only owing to increasing time complexity of the network. The whole experiment took around 20 days to run despite the low values of employed bees and maximum number of iterations. Thus, there is a great need to minimize the time required to run the proposed algorithm which can be achieved either by introducing parallelism in the algorithm or by using the cluster of nodes in the network. Thus, the future work is to optimize the working of proposed algorithm on the spark cluster to minimize the time complexity and to improve the accuracy of the algorithm

REFERENCES

1. S. Mittal and O.P. Sangwan, "Metaheuristic Based Approach To Regression Testing," International Journal of Computer Science and Information Technologies, Vol. 6, No. 3, 2015.
2. S. Mittal and O.P. Sangwan, "Big Data Analytics using Machine Learning Techniques," 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 10-11 Jan. 2019.
3. S. Mittal and O.P. Sangwan, "Big Data Analytics Using Data Mining Techniques: A Survey," in book: Advanced Informatics for Computing Research, 264-273, 2nd International Conference, ICAICR, Shimla, 14-15 July 2018.
4. S. Mittal and O.P. Sangwan, "Big Data Analytics using Deep LSTM Networks: A Case Study for Weather Prediction," Advances in Science, Technology and Engineering Systems Journal, Vol. 4, No. 2, 2020.
5. H. Nimbark, R. Sukhadia and P.P. Kotak, "Optimizing Architectural Properties of Artificial Neural Network using Proposed Artificial Bee Colony Algorithm," IEEE, International Conference on Advances in Computing, Communications and Informatics, 24-27 Sep. 2014.
6. B.A. Garro and R.A. Vazquez, "Designing Artificial Neural Networks using Particle Swarm Optimization Algorithms," Hindawi, Computational Intelligence and Neuroscience, 2015.
7. X. Cai, N. Zhang, G. K.Venayagamoorthy and D. C. Wunsch, "Time Series Prediction with Recurrent Neural Networks Using a Hybrid PSO-EA Algorithm," IEEE, International Joint Conference on Neural Networks, Vol. 2, Aug. 2004.
8. A.K. Tripathy, S. Mohapatra, S. Beura and G. Pradhan, "Weather Forecasting using ANN and PSO," International Journal of Scientific & Engineering Research, Vol. 2, No. 7, 2011.
9. N.M. Nawi, A. Khan, M. Z. Rehman, H. Chiroma and T. Herawan, "Weight Optimization in Recurrent Neural Networks with Hybrid Metaheuristic Cuckoo Search Techniques for Data Classification," Hindawi, Mathematical Problems in Engineering, Vol. 4, Sep. 2015.
10. H.M. Albeahdili, T. Han and N.E. Islam, "Hybrid Algorithm for the Optimization of Training Convolutional Neural Network," International Journal of Advanced Computer Science and Applications, Vol. 6, No. 10, 2015.
11. Q. Kang, W.K. Liao, A. Agrawal and A. Choudhary, "A Hybrid Training Algorithm for Recurrent Neural Network using Particle Swarm Optimization-based Preprocessing and Temporal Error Aggregation," IEEE, International Conference on Data Mining Workshops, pp 812-817, 15 Dec. 2017.
12. T. Li, B. Wang, M. Zhou and J. Watada, "Short-term load forecasting using optimized LSTM networks based on EMD," Electrical Engineering and Systems Science, Signal Processing, Aug. 2018.
13. I. Salman, O.N. Ucan, O. Bayat and K. Shaker, "Impact of Metaheuristic Iteration on Artificial Neural Network Structure in Medical Data," Processes, MDPI, Vol. 6, No. 5, 2018.
14. B. Nakisa, M.N. Rastgoo, A. Rakotonirainy, F. Maire and V. Chandran, "Long Short Term Memory Hyperparameter Optimization for a Neural Network Based Emotion Recognition Framework," IEEE Access, Vol. 6, 2018.
15. H. Chiroma, U. A. Abdullahi, S. M. Abdulhamid, A. A. AlArood, L. A. Gabralla, N. Rana, L. Shuib, I. A. T. Hashem, D. E. Gbenga, A. I. Abubakar, A. M. Zeki and T. Herawan, "Progress on Artificial Neural Networks for Big Data Analytics: A Survey," IEEE Access, Vol. 7, 2018.
16. T.A. Rashid and S. Abdullah, "A Hybrid of Artificial Bee Colony, Genetic Algorithm, and Neural Network for Diabetic Mellitus Diagnosing," ARO-The Scientific Journal of Koya University, [S.I.], Vol. 6, No. 1, pp. 55-64, 2018.
17. O. Zavalnyi, G. Zhao, Y. Savchenko and W. Xiao, "Experimental Evaluation of Metaheuristic Optimization of Gradients as an Alternative to Back propagation," 4th International Conference on Computer and Communications, IEEE, 7-10 Dec. 2018.
18. J. Zhang and F.B. Gouza, "GADAM: Genetic-Evolutionary ADAM for Deep Neural Network Optimization," Computer Science, Machine Learning, Cornell University, 2018.
19. A.P. Piotrowski, M. Osuch, M.J. Napiorkowski, P.M. Rowinski and J.J. Napiorkowski, "Comparing Large Number of Metaheuristics for Artificial Neural Networks Training to Predict Water Temperature in a Natural River," Computers & Geosciences, Vol. 64, pp 136-151, 2014.
20. E. Nourani, A.M. Rahmani and A.H. Navin, "Forecasting Stock Prices using a hybrid Artificial Bee Colony based Neural Network," International Conference on Innovation, Management and Technology Research (ICIMTR2012), Malacca, Malaysia, 21-22 May 2012.
21. J.M. Ordonez, C. Cobos, M. Mendoza, E.H. Viedma, F. Herrera and S. Tabik, "Framework for the Training of Deep Neural Networks in TensorFlow using Metaheuristics," International Conference on Intelligent Data Engineering and Automated Learning, pp 801-811, 2018.
22. X.Y.Chen, K.W.Chau and A.O.Busari, "A Comparative Study of Population-Based Optimization Algorithms for Downstream River Flow Forecasting by a Hybrid Neural Network Model," Engineering Applications of Artificial Intelligence, Vol. 46, pp 258-268, 2015.
23. <https://www.kaggle.com/PROPPG-PPG/hourly-weather-surface-brazil-southeast-region>.

AUTHORS PROFILE



Shweta Mittal received the B.Tech. degree in Computer Science and Engineering from B.R.C.M. College of Engineering & Technology, Bahal, Haryana, India and M.Tech. degree from Guru Jambheshwar University of Science and Technology (GJU S&T), Hisar, Haryana, India in 2013 and 2015, respectively. She is currently pursuing the Ph.D. degree in Computer Science and Engineering from Guru Jambheshwar University of Science and Technology (GJU S&T), Hisar, Haryana, India. Her current research interests are Big Data Analytics, Data Mining, Deep Learning, Metaheuristic Techniques and Regression Testing.



Om Prakash Sangwan received the M.Tech. degree (Hons.) in Computer Science and Engineering and the Ph.D. degree in Computer Science and Engineering from the Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India, where he is currently a Professor with the Department of Computer Science and Engineering. He was the Deputy Director with the Amity Resource Centre for Information Technology and Head, Cisco Regional Networking Academy, Amity Institute of Information Technology, Amity University, Uttar Pradesh. He was an Assistant Professor with the Department of Computer Science and Engineering, School of Information and Communication Technology, Gautam Buddha University, Greater Noida, Uttar Pradesh. He is also a CISCO Certified Network Associate (CCNA) and a CISCO Certified Academic Instructor (CCAI). He has number of publications in international and national journals and conferences. His research interests include software engineering focusing on planning, designing, testing, metrics, and application of neural networks, and fuzzy logic and neuro-fuzzy. He is a member of Computer Science Teacher Association (CSTA), New York, USA, International Association of Engineer (IAENG), Hong Kong, and International Association of Computer Science and Information Technology (IACIST), USA, a professional member of Association of Computing Machinery, IEEE, and a Life Member of Computer Society of India.