JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 36, 231-241 (2020)
DOI: 10.6688/JISE.202003_36(2).0004

Evaluation of Genetic Algorithm Optimization in Machine Learning

DHEYAA SHAHEED AL-AZZAWI
Department of Computer
College of Computer Science and Information Technology
Wasit University
Wasit, 52001 Iraq
E-mail: dalazzawi@uowasit.edu.iq

The manuscript is presenting the usage scenarios of Genetic Algorithm is one of the high performance algorithms for the engineering optimization. The aim of this paper is to present the effectual implementation with the prominent evolutionary computation method of genetic algorithm for the data analytics and in specific usage towards the multiprocessor scheduling approach. The presented algorithm is giving the superior outcomes on the assorted parameters so that the concrete and tangible results from genetic algorithm can be presented. The traditional or previous methods were making use of greedy based approach and thereby the results were not effective that motivated and prompted to work on the genetic algorithm based optimization approach. The implementation scenarios are simulated in MATLAB as well as Java based development libraries and found the outcome quite effective and valuable in terms of optimization.

Keywords: genetic algorithm, engineering optimization, search based optimization, data analysis, artificial intelligence

1. INTRODUCTION

The space of information revelation and prescient investigation is progressively engaged and subordinate towards machine learning and profound learning based applications. Tremendous calculations and approaches are accessible in machine learning for logical applications and answers for genuine issues [1, 2]. Comprehensively, there are three sorts of methodologies in the machine learning which are broadly coordinated for the critical thinking and prescient mining. These methodologies incorporate regulated learning, unsupervised learning and fortified learning. These methodologies are utilized according to the particular space of execution and exactness required [3-5].

The key problem in the previous approaches exist in terms of global optimization which are required to be addressed and solved using soft computing and evolutionary algorithms which provides the higher degree of accuracy and minimum cost factor. In the greedy based or classical heuristic based approaches there are lacking on the parameters of optimization which can be solved using evolutionary and population based approaches to have the outcomes on multiple parameters [6-8].

The previous works on the greedy algorithms were not quite effective and used the heuristics to have the nearby solutions which can be solved and optimized. In addition, the previous approaches make use of techniques which can be further enhanced using

Received September 17, 2019; revised September 17&26, 2019; accepted October 2, 2019.
Communicated by Osamah Ibrahim Khalaf.

soft computing and nature inspired approaches because nature provides the algorithms and techniques which are quite optimized, see Fig. 1 [9, 10].

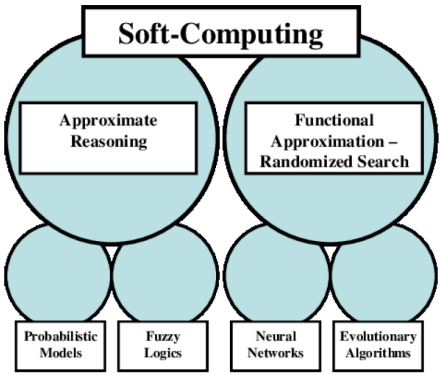

Fig. 1. Optimization approaches.

It is a logic of multi-valued which permits moderate values to be expressed between traditional evaluations like true/false. It provides a simple manner to reach at specific conclusion depend at the imprecise, ambiguous, vague, noisy, or missing input information [11-13].

Neural network is a collection of processing units named neurons, communicate by sending signals to each other through a huge number of weighted connections network depend on biological neurons. Machine learning focus on the changes of the systems when implement tasks related with artificial intelligence mechanism. The changes may be in either existing system or newly proposed system [14-17].

The domain of engineering optimization and search based data analytics is huge and needs utmost care in assorted mains. There are number of approaches and algorithms, which are used for the search, based optimization to solve the real time problems. Machine Learning, Deep Learning and Predictive Analytics are the key areas of research in grouped spaces of usage including designing, account, financial matters, continuous imaging and numerous others. The specialists are taking a shot at various devices and advances including open source and possess created systems so the higher level of exactness can be accomplished.

Prominent Algorithms for Optimization includes Spiral Dynamic Algorithm (SDA), Plant Propagation Algorithm, MBO: Artificial Chemistry Reacting Optimizer Algorithm, Spider Monkey Optimizer (SMO) algorithm, Water Cycle Algorithm , Elephent herding optimization (EHO), Artificial Eco Algorithm (AEA), Bird Mating Optimizer, Artificial immune systems (AIS) , Marriage in Honey Bees Optimization, Vortex Search Algorithm, Mine blast algorithm, Grey Wolf Optimizer, Cultural Algorithms (CA), Simulation annealing (SA), Ant Colony Optimizer (ACO), The Raven Roosting Optimization Algorithm, Black Holes Algorithm, Chemical Reaction Algorithm, Cuckoo Search (CS), Differential Search Algorithm (DS, Bull optimization algorithm, Bees Algorithm, Cuttlefish Algorithm, Glowworm Swarm Optimization (GSO), Intelligent Water Drops algorithm, Boids, Bacterial Foraging Optimization, Altruism Algorithm, Golden Ball, Bacterial Evolutionary Algorithm (BEA), (GSA), [18-21].

2. LITERATURE REVIEW

H. Duan and P. Qiao [4], designed and implemented a new algorithm named Pigeon-Inspiration-Optimizer (PIO), they used this algorithm to solve the problems of robot walk planning, the empirical results were compared with the values of other traditional old results works and show the high performance of the proposed manner.

S. Oreski and G. Oreski [5], proposed a new expert elicitation algorithm. It consists of mixing genetic algorithm within the neural net, it plays the main role in selecting the best set of the features for enforcing the recognition performance in the credit bank, experimentally, show that this proposed algorithm presents the high performance of recognition comparing with other elicitation algorithms.

P. Ghamisi *et al.* [9], proposed a new approach for choosing the features from the objects, this approach depend at mixing a genetic algorithm within particle-swarming algorithm, then the fitness value constrained mainly at the support vector machine. This new approach was implemented with two projects. The first one is Indian wood dataset and the results show that this algorithm present results with high accuracy of classification, and the second project, was implemented to distinguish between the objects walking at the roads and the background, and experimentally, the results proved the high performance of the proposed approach in the features selection process.

D. S. Sabr [12], which mixing the points of descriptors within corner detection algorithm, and resulted with new approach which was experimentally proved that this new approach gave good results and high performance in the detection and the recognition objects between multiple objects in side occluded pattern. The main light point of this approach, the detection and recognition process didn't need and training process.

C. Yoo *et al.* [13], produce a new bioinformatics manner and novel statistical learning, which were used for training the very big data set of medicine database with its relationships, this proposed manner was experimentally proven its good performance, compared with other methods and this would open new horizon to use a novel statistical learning to understand the human disease.

C. Li *et al.* [17], used a deep learning for detecting the fault in the gearbox and bearing systems in the machine, by putting sensor inside the machine and collect the vibration signals that produced from the gearbox and bearing system, these signals received in three types, time, frequency and time-frequency domains. All three types of value were used as a feature for the learning, the tools of learning were produced named Gaussian Bernoulli Deep Boltzmann Machine (GDBM), and the suggested manner were experimentally compared with old traditional such SVM and GRBM and the results show the performance and more than 95% classification performance.

3. STRUCTURE OF THE PRESENT WORK

Genetic algorithm is an optimization technique and heuristic search that mimic the task of natural growth and is depicted in Fig. 2, it simulates growth of humanity by the assists of natural selection and choice the best element and discards the rest part [22-24].

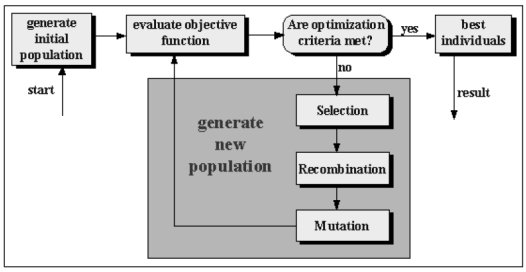

Fig. 2. Genetic algorithm.

Table 1. Individuals in GA.

Individual 1	12	225	5
Individual 2	321	4	34

Table 2. Samples in GA.

Sample 1	2	2	1
Sample 2	1	2	1

Table 3. Offspring in GA.

Offspring 1	321	4	5
Offspring 2	12	4	5

$$Offspring = {}^{a}(Parent_2 - Parent_1) + Parent_1$$

Tables 1-3 present the Individuals, Samples and Offspring respectively, which are processes in the GA based approach. Consider again the following two individuals as in Table 4 with 3 variables each [25-28]:

α was selected here as follow in Table 5, then the new productions of individuals are figured as follow in Table 6.

Table 4. Processing of parameters.

Individual 1	12	225	4
Individual 2	321	4	34

Table 5. Selection process.

Sample 1	0.5	1.3	–0.1
Sample 2	0.2	0.9	0.5

Table 6. New individual processing.

Offspring 1	67.5	1.9	2.1
Individual 2	23.1	8.2	19.5

The approach of genetic algorithm is used in the assorted engineering domains [29-31] for the best fit search based optimization. Following are the key components in the GA based approach as depicted in Fig. 3.

- Natural Selection
- Population Generation
- Fitness Function
- Selection Process
- Crossover
- Mutation Process
- Termination

- Fitness Selection
- Optimization
- Best Fit Function and Selection of Candidate
- Final Optimization and Search Process

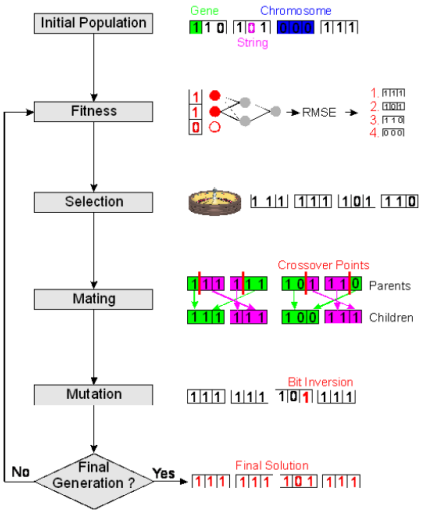

Fig. 3. Flowchart of genetic algorithm.

4. SIMULATION RESULTS AND OUTCOME

The simulation of the GA can be done in many tools and this process is widely integrated for the getting of results and simulation outcomes [32-35]. This work is having the simulation patterns in MATLAB and Java with the comparison to other various approaches with GA. The running phases are presented in Fig. 4, with different parameters of GA with the optimization perspectives in Fig. 5.

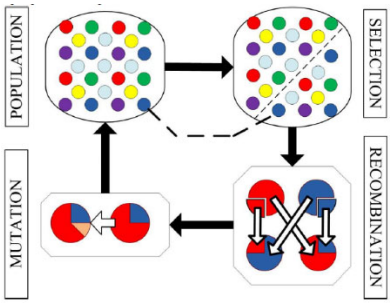

Fig. 4. GA automation process.

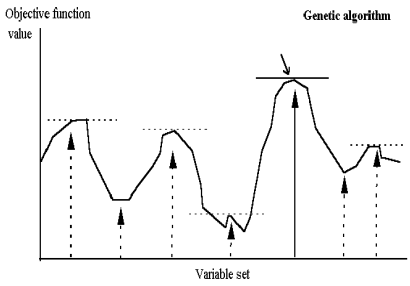

Fig. 5. GA Thematic function process.

The genetic algorithm optimization function is used to pick the fitness out and the overall probabilities of the best results [36-38].

5. SIMULATION OUTCOME ON PRODUCTION ENGINEERING DEFECTS PREDICTION

There are overall set of 1000 records in the dataset and following is the sample sheet, as depicted in Fig. 6. The datasets are simulated with the execution run on the production engineering dataset to extract the defects in the production line as shown in Fig. 7.

```
--------------------------------------------------
Reading DataSet for Deep Analysis and Optimization
--------------------------------------------------

A =

    1.0000   10.0000    1.0000         0    0.0370
    3.0000    9.0000    6.0000         0    0.2221
    3.0000    2.0000    9.0000    2.0000    2.3338
    4.0000    6.0000    3.0000    2.0000    2.1116
    6.0000    7.0000    9.0000    2.0000    2.3353
    8.0000    8.0000    4.0000    2.0000    2.1495
   10.0000    9.0000    7.0000         0    0.2621
   11.0000    8.0000    8.0000    2.0000    2.3003
   12.0000    5.0000    3.0000    2.0000    2.1130
   12.0000    4.0000    1.0000    2.0000    2.0377
   13.0000    8.0000    5.0000         0    0.1885
   15.0000    1.0000   10.0000    2.0000    2.3778
   16.0000    8.0000    1.0000    2.0000    2.0379
   17.0000    4.0000    6.0000    2.0000    2.2276
   17.0000   10.0000    7.0000    2.0000    2.2662
   18.0000   10.0000    3.0000    1.0000    1.1144
   18.0000    8.0000    3.0000         0    0.1145
```

Fig. 6. Sample sheet of 1000 records in the dataset.

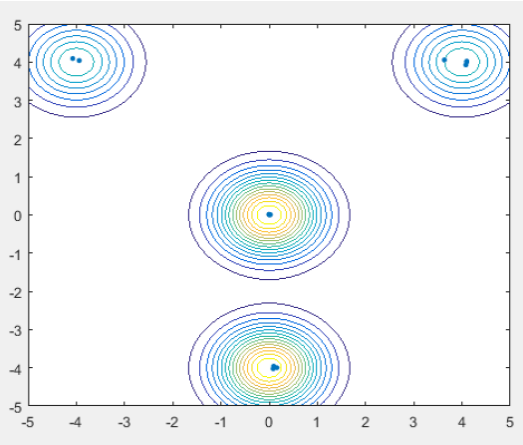

Fig. 7. Simulation process.

In each attribute, the value of machine defect is presented and then passed to the Genetic Algorithm for getting the root cause; Table 7 shows the comparison between the results of outcomes.

Table 7. Comparison of outcomes.

Simulation Attempt	GA	Honeybee	Firefly
1	94	86	78
2	90	85	82
3	99	86	86
4	96	85	79
5	96	80	78
6	91	84	79
7	99	87	84
8	90	80	80
9	91	84	83
10	96	81	82
11	96	81	81
12	96	84	85

Fig. 8 has the depiction of the outcomes on the performance level of traditional and presented approach on the assorted execution scenarios.

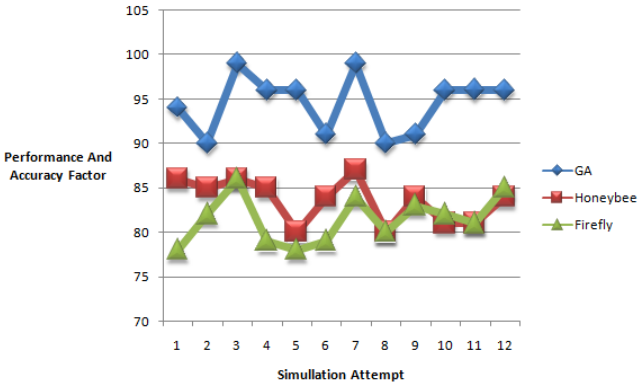

Fig. 8. Outcome and performance aspects.

6. IMPLEMENTATION OF MULTIPRCESSOR SCHEDULER WITH GA

Fig. 9 presents the view of the nodes in GA with the instances of execution in the simulation. Fig. 10 depicts the outcome on the approach without GA, while Fig. 11 depicts the outcome with GA based approach. It is found from the outcomes that the GA based approach is quite effective.

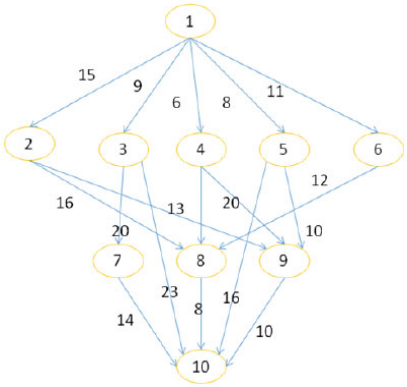

Fig. 9. View of multiprocessor optimization in GA.

P1

T1	T2
0-6	6-18

P2

NO PROCESS	T5	T6	T7	T8
0-6	6-16	16-28	28-42	42-50

P3

NO PROCESS	T3	T4	T9	T10
0-6	6-26	26-46	46-56	56-64

Execution Time: 64

Fig. 10. Implementation outcomes without GA.

P1

T1	T4	T5	T9	T10
0-6	6-26	26-36	36-46	46-54

P2

NO PROCESS	T3	T7
0-6	6-26	26-40

P3

NO PROCESS	T2	T6	T8
0-6	6-19	19-31	31-39

Execution Time: 54

Fig. 11. Implementation outcomes with GA.

Multi Processor Scheduler without Genetic Algorithm is thusly introduced as the initial step of the procedure, Select the undertakings which are having no parent from the information task rundown and run them in parallel among three processors. In the event that the no. of assignment chose is under 3 Run the chose procedure in any of the processor and end the staying till the execution finishes. In the event that there are numerous conceivable approaches to run the procedure, select the route by which the procedure can be executed quicker. As a second step in booking, select the assignment which has its parent executed completely. These assignments are planned for procedures in such a design, that there is no time slack by which a most extreme effectiveness can be accomplished from every processor. On finish of each errand its fulfillment status is set to 1 so that, before choosing an assignment for execution itself the parent consummation is checked. In the event that an errand has at least 2 parents, before executing the undertaking the majority of its parent finish are checked and ensure that the assignment is executed simply after the most extreme burst time returned by any of its parent. Burst time returned by the last errand executed is considered as the all out execution time of the absolute procedure.

Multi Processor Scheduler with Genetic Algorithm incorporates the accompanying the execution strategy is comparable as above aside from the way that the errand chose for execution is returned by Genetic algorithm which is produced for the equivalent. GA is a customary developmental algorithm which has a gathering of people picked for the technique. We have populace estimate and no. of ages as approvals to restrain the child components in transit they are. GA has 2 principle tasks to be specific traverse and change. From the arrangement of people parents are arbitrarily picked and a traverse is connected over them in which they bit esteems are changed with the goal that another child procedure is found. Similarly, numerous child things are created by the haphazardly picked individual parents. The child to go about as parent in cutting edge should pass the base fitness esteem and ought not to surpass the extreme no. of ages that are permitted to perform. On the off chance that a child with worthy fitness is accomplished before the most extreme age we stop the execution and return the qualities.

7. CONCLUSIONS

The algorithms of machine learning, deep learning, data science and knowledge discovery are closely associated domains for the scientific and engineering applications. These algorithms can be used for the development of new algorithms and solving the engineering optimization problems in different domains for the social as well as scientific domains. These algorithms are having custom functions, which can be updated as per the requirements of the dynamic datasets to achieve and higher degree of accuracy and performance with related dimensions of effectiveness. The work is presenting the usage of Genetic Algorithm for the optimization in the engineering and related domains with the assorted scenarios. The proposed work can be further elevated using the nature inspired and hybrid soft computing approaches with the following Fuzzy Logic based integrations including Support Vector Machines, Swarm Intelligence, Meta-heuristics, Bayesian Network, Evolutionary Approaches and many others so that further the outcomes can be elevated to higher degree of accuracy and performance on the assorted parameters.

REFERENCES

1. K. Slavakis, G. B. Giannakis, and G. Mateos, "Modeling and optimization for big data analytics: (statistical) learning tools for our era of data deluge," *IEEE Signal Processing Magazine*, Vol. 31, 2014, pp. 18-31.
2. W. V. D. Aalst, "Data science in action," *Process Mining*, Springer, 2016, pp. 3-23.
3. D. S. Al-Azzawy, "Application of Haar-like features in three AdaBoost algorithms for face detection," *Journal of Wassit for Science & Medicine*, Vol. 4, 2011, pp. 38-54.
4. H. Duan and P. Qiao, "Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning," *International Journal of Intelligent Computing and Cybernetics*, Vol. 7, 2014, pp. 24-37.
5. S. Oreski and G. Oreski, "Genetic algorithm-based heuristic for feature selection in credit risk assessment," *Expert Systems with Applications*, Vol. 41, 2014, pp. 2052-2064.
6. D. S. Al-Azzawy and F. M. Al-Rufaye, "Arabic word clustering using K-means algorithm," in *Proceedings of Annual Conference on New Trends in Information and Communications Technology Applications*, 2017, pp. 263-267.
7. D. S. Al-Azzawi, "Recognition and diagnosis perspectives for cancer patterns in the medical digital images," *Journal of Advanced Research in Dynamical and Control Systems*, Vol. 11, 2019, pp. 1090-1098.
8. D. S. Al-Azzwy, "Eigenface and SIFT for gender classification," *Journal of Wassit for Science & Medicine*, Vol. 5, 2012, pp. 60-76.
9. P. Ghamisi and J. A. Benediktsson, "Feature selection based on hybridization of genetic algorithm and particle swarm optimization," *IEEE Geoscience and Remote Sensing Letters*, Vol. 12, 2015, pp. 309-313.
10. M. Thakur, S. S. Meghwani, and H. Jalota, "A modified real coded genetic algorithm for constrained optimization," *Applied Mathematics and Computation*, Vol. 235, 2014, pp. 292-317.
11. T. Bäck, D. B. Fogel, and Z. Michalewicz, eds., *Evolutionary Computation 1: Basic Algorithms and Operators*, CRC Press, FL, 2018.
12. D. S. Sabr, "Points descriptor in pattern recognition: A new approach," *Iraqi Journal for Computer and Informatics*, Vol. 43, 2017, pp. 292-317.
13. C. Yoo, L. Ramirez, and J. Liuzzi, "Big data analysis using modern statistical and machine learning methods in medicine," *International Neurourology Journal*, Vol. 18, 2014, p. 50.
14. T. Santhanam and M. S. Padmavathi, "Application of K-means and genetic algorithms for dimension reduction by integrating SVM for diabetes diagnosis," *Procedia Computer Science*, Vol. 47, 2015, pp. 76-83.
15. J. Ashraf and S. Latif, "Handling intrusion and DDoS attacks in software defined networks using machine learning techniques," in *Proceedings of IEEE National Software Engineering Conference*, 2014, pp. 55-60.
16. D. S. Al-Azzawy, "Review on challenges in building knowledge based societies through e-governance framework: Technology and management issues from global perspective," *Indian Journal of Science and Technology*, Vol. 10, 2017.
17. C. Li, R. V. Sánchez, G. Zurita, M. Cerrada, and D. Cabrera, "Fault diagnosis for

rotating machinery using vibration measurement deep statistical feature learning," *Sensors*, Vol. 16, 2016, p. 895.
18. B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms," *Expert Systems with Applications*, Vol. 41, 2014, pp. 1476-1482.
19. S. Salcedo-Sanz, A. Pastor-Sánchez, L. Prieto, A. Blanco-Aguilera, and R. García-Herrera, "Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization-extreme learning machine approach," *Energy Conversion and Management*, Vol. 87, 2014, pp. 10-18.
20. O. Kisi, J. Shiri, S. Karimi, S. Shamshirband, S. Motamedi, D. Petković, and R. Hashim, "A survey of water level fluctuation predicting in Urmia lake using support vector machine with firefly algorithm," *Applied Mathematics and Computation*, Vol. 270, 2015, pp. 731-743.
21. N. Izadyar, H. Ghadamian, H. C. Ong, C. W. Tong, and S. Shamshirband, "Appraisal of the support vector machine to forecast residential heating demand for the district heating system based on the monthly overall natural gas consumption," *Energy*, Vol. 93, 2015, pp. 1558-1567.
22. A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques," *International Journal of System Assurance Engineering and Management*, Vol. 8, 2017, pp. 937-953.
23. X. S. Yang, *Nature-Inspired Optimization Algorithms*, 2014, Elsevier.
24. J. S. Chou, M. Y. Cheng, Y. W. Wu, and A. D. Pham, "Optimizing parameters of support vector machine using fast messy genetic algorithm for dispute classification," *Expert Systems with Applications*, Vol. 41, 2014, pp. 3955-3964.
25. L. C. Padierna, M. Carpio, A. Rojas, H. Puga, R. Baltazar, and H. Fraire, "Hyper-parameter tuning for support vector machines by estimation of distribution algorithms," *Nature-Inspired Design of Hybrid Intelligent Systems*, Springer, Cham, 2017, pp. 787-800.
26. E. Tuba and Z. Stanimirovic, "Elephant herding optimization algorithm for support vector machine parameters tuning," in *Proceedings of IEEE 9th International Conference on Electronics*, *Computers and Artificial Intelligence*, 2017, pp. 1-4.
27. D. S. Al-Azzawy and S. A. Diwan, "Design of intelligent agent based management security system for e-government," *Journal of Al-Qadisiyah for Computer Science and Mathematics*, Vol. 9, 2012, pp. 131-142.
28. C. Li, S. Li, and Y. Liu, "Least squares support vector machine model optimized by moth-flame optimization algorithm for annual power load forecasting," *Applied Intelligence*, Vol. 45, 2016, pp. 1166-1178.
29. H. Faris, M. A. Hassonah, A. Z. Ala'M, S. Mirjalili, and I. Aljarah, "A multi-verse optimizer approach for feature selection and optimizing SVM parameters based on a robust system architecture," *Neural Computing and Applications*, Vol. 30, 2018, pp. 2355-2369.
30. L. Olatomiwa, S. Mekhilef, S. Shamshirband, K. Mohammadi, D. Petković, and C. Sudheer, "A support vector machine-firefly algorithm-based model for global solar radiation prediction," *Solar Energy*, Vol. 115, 2015, pp. 632-644.
31. F. Kang, J. S. Li, and J. J. Li, "System reliability analysis of slopes using least squares support vector machines with particle swarm optimization," *Neurocompu-*

ting, Vol. 209, 2016, pp. 46-56.
32. A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, Vol. 18, 2016, pp. 1153-1176.
33. T. Kanimozhi and K. Latha, "An integrated approach to region based image retrieval using firefly algorithm and support vector machine," *Neurocomputing*, Vol. 151, 2015, pp. 1099-1111.
34. G. I. Sayed, A. Darwish, A. E. Hassanien, and J. S. Pan, "Breast cancer diagnosis approach based on meta-heuristic optimization algorithm inspired by the bubble-net hunting strategy of whales," in *Proceedings of International Conference on Genetic and Evolutionary Computing*, 2016, pp. 306-313.
35. D. S. Al-Azzawi, "Cloud robotic: An effectual analysis with its cavernous perspectives," *Journal of Advanced Research in Dynamical and Control Systems*, Vol. 11, 2019, pp. 1104-1109.
36. A. Tharwat, A. E. Hassanien, and B. E. Elnaghi, "A BA-based algorithm for parameter optimization of support vector machine," *Pattern Recognition Letters*, Vol. 93, 2017, pp. 13-22.
37. R. Khatibi, M. A. Ghorbani, and F. A. Pourhosseini, "Stream flow predictions using nature-inspired firefly algorithms and a multiple model strategy-directions of innovation towards next generation practices," *Advanced Engineering Informatics*, Vol. 34, 2017, pp. 80-89.
38. X. Zhang, W. Chen, B. Wang, and X. Chen, "Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization," *Neurocomputing*, Vol. 167, 2015, pp. 260-279.

Dheyaa Shaheed Al-Azzawi was born in Wasit Province, Iraq in 1972. He received the B.S. degree from the University of Babylon, College of Sciences, Babylon, Iraq in 1996 and the M.S. degree from the University of Babylon, College of Sciences, Babylon, Iraq in 2001. The Ph.D. degree from University of Information Technology and Communications, Baghdad, Iraq 2007, all in Computer Sciences, he is currently a Professor at Wasit University, College of Computer Sciences and Information Technology. His research interests include artificial intelligence, computer vision, and object recognition.