Evaluating the Effect of Optimized Voting Using Hybrid Particle Swarm and Grey Wolf Algorithm on the Classification of the Zoo Dataset

Doaa S. Khafaga[*1], Hussein Alkattan[2], Alhumaima A. Subhi[3]

[1] Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

[2]Department of System Programming South Ural State University; pr. Lenina 76, Chelyabinsk, 454080 Russia

[3]Electronic and Computer Center, University of Diyala, Baqubah MJJ2+R9G, Iraq

Emails:dskhafga@pnu.edu.sa; alkattan.hussein92@gmail.com, alhumaimaali@uodiyala.edu.iq

Abstract

When there are numerous possible solutions for a given class in a given problem, majority voting or plurality voting is typically employed. One common technique for improving classification accuracy is bagging, which involves training many classifiers on slightly different datasets and then voting on the combined results. In this research, we examine how alternative voting procedures affect the efficiency of two distinct classification algorithms applied to datasets of varying complexity. Despite the increased computing cost associated with determining preference order, the results show that the single transferable vote can be a suitable alternative to plurality voting.

Keywords: Zoo data; Voting classifier; Support vector manchines; Neural networks; Decision trees.

1. Introduction

Since the rise of the Internet and the usage of increasingly massive databases, data processing techniques have been the primary focus of study in many disciplines. Without the aid of specialized instruments or automated software procedures, it is challenging for a human user to grasp or manage such data. Many applications, including categorization, recommendation systems, pattern recognition, etc., rely heavily on machine learning to meet these requirements. Ensemble techniques have been validated as an effective strategy for enhancing classification precision. Bagging, boosting, and stacking are a few of the most common strategies. Most existing bagging-based classification algorithms average the results from separate voting on each ensemble component. Despite its apparent ease, this method may not always be the most effective when dealing with multiclass issues. In this research, we employ ensembles of k-nearest neighbor and Nave Bayes classifiers aggregated by four different voting systems and evaluate their performance in terms of classification accuracy to investigate the impact of different voting approaches paired with bagging.

The authors' algorithms employed here were all written, with no external libraries utilized. To better present our paper's arguments, we've laid them down as follows. The section on related work is presented in Section II, a brief description of bagging and the voting techniques under investigation

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022

are presented in Section III, and the two base classifiers that were used to evaluate the effectiveness of bagging are shown in Section IV. In Section V, we show three case studies illustrating the use of the presented approaches to solving multi-class classification problems of progressively increasing difficulty. In Section VI, we draw some conclusions and offer some suggestions for future research. The structure of this paper is shown in Figure 1.

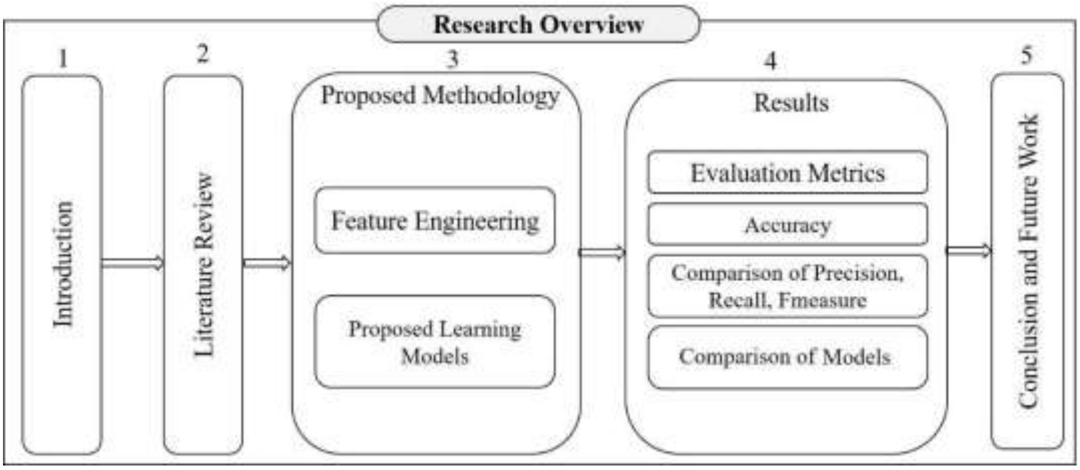

Figure 1: Research Paper Overview

2. Related Work

The field of machine learning that deals with classifying data have seen the most success thus far. Consequently, the perpetual effort is put into developing new methods that improve categorization precision. To this end, one approach is to employ many classifiers working together. Combining the results of many classifiers into a single, more precise prediction, bagging [1], and boosting [2] established the first compelling ensemble approaches. But the traditional form employs the majority or plurality vote to sum together the results of separate classifiers. To this end, researchers have been probing the intersection of voting theory and ensemble techniques, with some preliminary findings provided below.

In [3], we see a variety of voting algorithms that may be used to evaluate ensembles of classifiers that have been trained using the bagging approach. MLPs are used as classifiers (MLPs). The training and testing datasets are split 50/50 between numeric and uppercase characters. Datasets were processed by a variety of MLP architectures (small and large). The findings indicated that the full rule, the product rule, and the Borda count are the most influential voting mechanisms for small MLPs. Additionally, huge MLPs benefited from Borda count's effective recognition. Reconciliation modeling, a technique for integrating classification models, is used in [4], along with bagging, to evaluate the effectiveness of various voting techniques. Classification trees were chosen as the classifier of choice, with the dataset split in two (one for training and one for testing). Bagging's generalization error performance under several voting schemes is demonstrated experimentally, including plurality, anti-plurality, Borda count, a plurality with elimination, and Condorcet's pairwise comparison.

The test data came from various sources and configurations, including noisy and noiseless datasets of varying sizes, datasets with and without labels, and classes with and without labels. Performances of anti-pluralism and plurality were stronger when there were only two classes. The Borda count outperformed competing voting systems in tests conducted in noisy environments. According to the results, the classification accuracy of the various voting techniques utilized is as follows: Borda count, Condorcet's pairwise comparison, anti-plurality, plurality, and plurality with elimination. Support, strength, and democratic voting were used with bagging to the MLEM2 rule induction system [5]. One hundred samples of data were used to create a bootstrapped sample of the LERS (Learning from Examples based on Rough Sets) classifier. The LERS classifiers were fed rules

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022

generated by using the MLEM2 induction rule to each bootstrap set. The trials included 16 datasets, and the error rates for each rule set used in bagging were reported. This had the error rates for the initial forecast made without aggregation. When a particular threshold number of classifiers was utilized, the error rate stabilized, demonstrating the bagging's stability (e.g., for the breast cancer dataset, the error rate became constant at an ensemble size of 27 with strength rule voting). Overall, the results were best for the democratic voting method and the worst for the voting method based on strength.

Bagging's effects on text categorization are detailed in [6]. The effectiveness of bagging was evaluated against the gold standard classifier after binary decision trees were generated using the C4.5 method. The tests were run on two datasets (the Reuter's and the Markiza collections) with a maximum of 200 classifiers per dataset. It was clear that both bagging and the traditional decision tree classifier performed accurately. The findings of the bagging approach were superior to those of the conventional method for more common classes and vice versa for less common ones. Bagging achieved superior accuracy when using more than 20 classifiers for the Reuter's collection and 10 classifiers for the Markiza collection. In [7], we examine the effectiveness of ensemble techniques on a subset (ten percent) of the KDD CUP '99 dataset (i.e., 494020 records). There are 41 attributes associated with data instances, and the labels that must be assigned are "normal" for the records exhibiting typical behavior and "attack" for malicious ones. Denial-of-service attacks, user-to-root attacks, and remote-to-local attacks are the three main types of malicious attacks. Bagging, AdaBoost, and the baseline C4.5 algorithm were all compared. AdaBoost outperformed Bagging and C4.5 by a small margin, with an error of 1.95 percent vs. 1.99 percent, respectively. In [8], a differential evolution-based weighted voting ensemble learning classifier (DEWVote) was introduced.

A differential weighted voting (DEWVote) classifier employed five different classifiers—C4.5, Naive Bayes, Bayesian Nets, k-Nearest Neighbour (kNN), and ZeroR—instead of just one, as a kind of diversity. This contrasts bagging, which uses the same base classifier every time. While DEWVote's voting mechanism is similar to majority voting, the outputs of the basic classifiers are given more weight in the final decision. Weights for each classifier were optimized using the differential evolution approach, with a more significant weight being given to classifiers that performed well on individual instances. Fifteen datasets were used to evaluate the four popular voting algorithms (bagging, AdaBoost, majority voting, and DEWVote). Naive Bayes was utilized as a classifier in both Bagging and AdaBoost. On 13 of the 15 datasets tested, DEWVote achieved the best results [9-13].

3. Proposed Model

Voting is a method agreed upon by most people as the best way to decide between competing options when making a decision. Each voter has the privilege of choosing the candidate(s) they like. After tallying the votes, it is possible to draw conclusions; in most cases, such findings will be in line with the people's will. To illustrate, if there are several solutions to a problem and you pick the worst one, you may have a poorer result than you would have gotten otherwise. The difficulty we are having is in determining how to categorize things. Numerous methods, including parametric and nonparametric ones and heuristic ones, those based on logic and probabilities, and so on, have been presented for building voting systems. The structure of the proposed model is shown in Figure 2.

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022

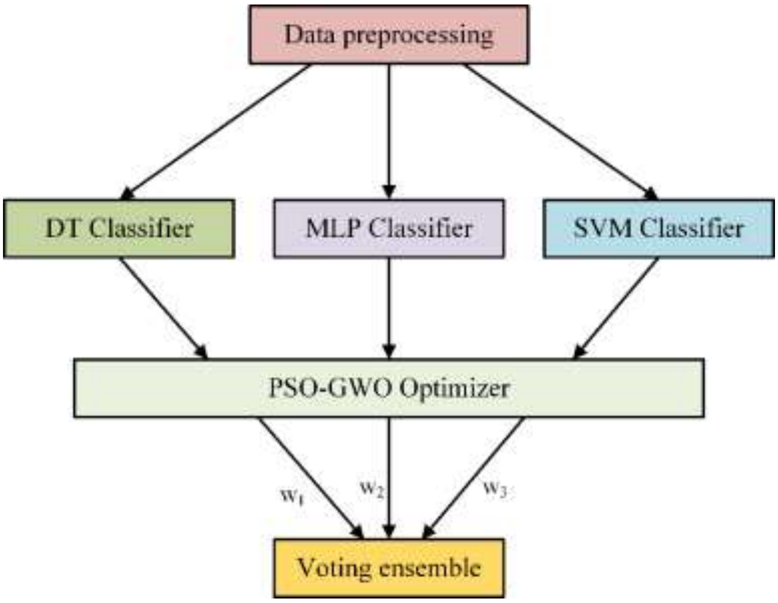

Figure 2: The structure of the proposed approach

Various classifiers, such as decision trees, kNN, multilayer perceptron, and so on, can be used in a voting system with different learning techniques, with the same training dataset for each. As a result, the predictions made by the various classifiers will vary [14-20]. A voting method where the victorious group represents a large proportion of the population can yield the most accurate forecast. The benefit of using many classifiers in an ensemble is that they are less likely to make the same error all. The usage of a single classifier across many training datasets is another option. In this study, we adopt a bagging strategy in which a single classifier is fed data from many datasets. One may also train the same classifiers on different feature samples in addition to bagging. The Random Forest method takes this concept one step further by combining the trained models of many decision trees that were created using various samples and independent random selection of their characteristics. The goal of the resampling method known as "bagging" (sometimes called "bootstrap aggregating"), which uses the original training dataset as input, is to generate many datasets that differ from the original in small ways. To illustrate, suppose we have a training dataset, D0, and we want m different classifiers. Therefore, we generate m additional datasets, Di, for I = 1,..., m. All classifiers, marked by Ci, are free to utilize the same method. To generate Di, we take independent samples from D0 according to a uniform sampling distribution with replacement. This means duplicate instances of the same set Di can be drawn from D0 several times. Ci is a classifier that is trained with each new group. Next, we utilize a voting method to average the results from all m classifiers. Each sample has a P1 = 1 - 1 / n probability of being left out of a single sampling, where n is the total number of instances in the D0 dataset. Pn = (1 - 1 / n) n is the chance that none of the examples in a dataset of size n will be chosen at random. Since each bootstrapped set Di contains around 63% separate samples from D0, this means that if the original dataset has a high number of occurrences, Di will have approximately 63% distinct samples. Compared to using a single trained classifier, the accuracy of classification using bagging significantly improves since the produced training sets Di to have less variation and, thus, less overfitting. Additionally, it performs better on noisy data. After that, we will go through a few different voting systems.

4. Simulation Results

This section presents the simulation results for the FIR LPF. The filter order (N) is 21. The sampling frequency (f_s) is 1Hz. The number of frequency samples is 512. Table 1 shows the simulation parameters for PSO and modified PSO, respectively. Both algorithms (PSO, and modified PSO) are run 50 times to obtain the best results.

Table 1: Classification results using the proposed method compared to other methods

	Accuracy	Sensitivity (TRP)	Specificity (TNP)	Pvalue (PPV)	Nvalue (NPV)	F-score

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022

NN	0.8511	0.9211	0.7692	0.8235	0.8929	0.8696
SVM	0.8562	0.9259	0.7692	0.8333	0.8929	0.8772
DT	0.8772	0.9434	0.7692	0.8696	0.8929	0.9050
PSO-GWO	0.9386	0.9843	0.8537	0.9259	0.9669	0.9542

The statistical analysis of the results achieved by the proposed approach is presented in Table 2. In this table, a comparison of the results achieved by the proposed method and neural network (NN), support vector machines (SVM), and decision trees (DT). These results show the effectiveness of the proposed methodology.

Table 2: Statistical analysis of the results recorded by the proposed method

	NN	SVM	DT	PSO-GWO
Number of values	10	10	10	10
Minimum	0.8411	0.8462	0.8672	0.9386
25% Percentile	0.8511	0.8562	0.8772	0.9386
Median	0.8511	0.8562	0.8772	0.9386
75% Percentile	0.8511	0.8562	0.8772	0.9386
Maximum	0.8611	0.8762	0.8872	0.9386
Range	0.02	0.03	0.02	0
10% Percentile	0.8421	0.8472	0.8682	0.9386
90% Percentile	0.8601	0.8742	0.8862	0.9386
95% CI of median				
Actual confidence level	97.85%	97.85%	97.85%	97.85%
Lower confidence limit	0.8511	0.8562	0.8772	0.9386
Upper confidence limit	0.8511	0.8562	0.8772	0.9386
Mean	0.8511	0.8572	0.8772	0.9386
Std. Deviation	0.004714	0.007379	0.004714	0
Std. Error of Mean	0.001491	0.002333	0.001491	0
Lower 95% CI of mean	0.8477	0.8519	0.8738	0.9386
Upper 95% CI of mean	0.8544	0.8624	0.8806	0.9386
Coefficient of variation	0.5539%	0.8608%	0.5374%	0.000%
Geometric mean	0.8511	0.8571	0.8772	0.9386
Geometric SD factor	1.006	1.009	1.005	1
Lower 95% CI of geo. mean	0.8477	0.8519	0.8738	0.9386
Upper 95% CI of geo. mean	0.8544	0.8624	0.8806	0.9386
Harmonic mean	0.851	0.8571	0.8772	0.9386
Lower 95% CI of harm. mean	0.8477	0.8519	0.8738	0.9386
Upper 95% CI of harm. mean	0.8544	0.8624	0.8806	0.9386
Quadratic mean	0.8511	0.8572	0.8772	0.9386
Lower 95% CI of quad. mean	0.8477	0.8519	0.8738	0.9386
Upper 95% CI of quad. mean	0.8544	0.8625	0.8806	0.9386
Skewness	0	1.908	8.83E-14	
Kurtosis	4.5	6.335	4.5	
Sum	8.511	8.572	8.772	9.386

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022

In addition, the Wilcoxon signed rank test is applied to evaluate the difference of the proposed methodology. The results of this test are presented in Table 3. These results prove the difference of the proposed method when compared to the other methods.

Table 3: Wilcoxon signed rank test of the recorded results of the proposed method

	NN	SVM	DT	PSO-GWO
Theoretical median	0	0	0	0
Actual median	0.8511	0.8562	0.8772	0.9386
Number of values	10	10	10	10
Wilcoxon Signed Rank Test				
Sum of signed ranks (W)	55	55	55	55
Sum of positive ranks	55	55	55	55
Sum of negative ranks	0	0	0	0
P value (two tailed)	0.002	0.002	0.002	0.002
Exact or estimate?	Exact	Exact	Exact	Exact
P value summary	**	**	**	**
Significant (alpha=0.05)?	Yes	Yes	Yes	Yes
How big is the discrepancy?				
Discrepancy	0.8511	0.8562	0.8772	0.9386

The accuracy of the proposed method compared to other methods is shown in Figure 3. The results of this figure confirm the effectiveness of the proposed method as it achieves the highest accuracy.

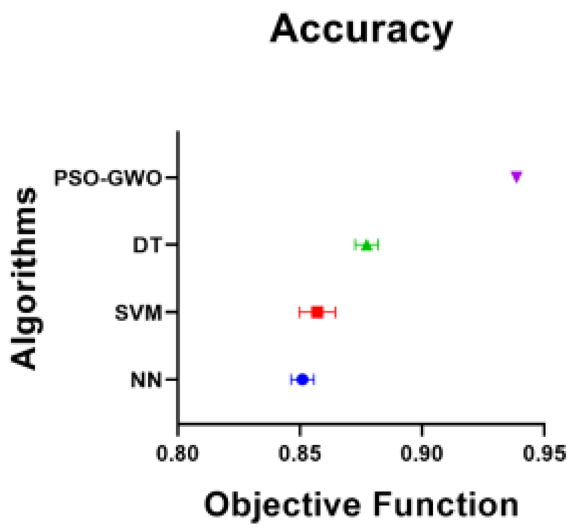

Figure 3: The accuracy of the proposed method compared to other methods

5. Conclusion

These tests demonstrate that the plurality technique is not always the best option, even though it is the simplest and most consistently effective voting system. The critical challenge is calculating the preference ranking, which is more time-consuming than just tallying votes. In certain circumstances, employing a particular technique readily lends itself to expressing this sorting, while in others, it is more challenging. The majority voting technique, typically used with ensemble-based categorization, is the plurality voting method; nevertheless, we may infer that the single transferable vote method can be a good option when feasible. Future research should naturally move toward using other classification algorithms and other classification tasks further to evaluate the impact of various voting techniques on bagging. In addition, a comprehensive statistical study

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022

of the performance of the voting techniques is required because of the limited size of the bootstrapped training sets.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] L. Breiman, "Bagging predictors", Machine Learning, vol.24, pp. 123- 140, 1996.
[2] R.E. Schapire, Y. Singer "Improved boosting algorithms using confidence-rated predictions", Machine Learning, vol. 37, pp. 297-336, 1999.
[3] M. Van Erp, L. Vuurpijl, L. Schomaker, "An overview and comparison of voting methods for pattern recognition", IWFHR '02 Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition, pp. 195-200, 2002.
[4] K.T. Leung, D.S. Parker, "Empirical comparisons of various voting methods in bagging", KDD '03 Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 595-600, 2003
[5] C. Cohagan, J.W. Grzymala-Busse, Z.S. Hippe, "A comparison of three voting methods for bagging with the MLEM2 algorithm", IDEAL'10 Proceedings of the 11th international conference on Intelligent data engineering and automated learning, pp. 118-125, 2010.
[6] K. Machová, F. Barčák, P. Bednár, "A bagging method using decision trees in the role of base classifiers", Acta Polytechnica Hungarica, vol.3, pp. 121-132, April 2006.
[7] N. Abdel Samee, E. M. El-Kenawy, G. Atteia, M. M. Jamjoom, A. Ibrahim et al., "Metaheuristic optimization through deep learning classification of covid-19 in chest x-ray images," Computers, Materials & Continua, vol. 73, no.2, pp. 4193–4210, 2022.
[8] A. A. Abdelhamid and S. R. Alotaibi, "Optimized two-level ensemble model for predicting the parameters of metamaterial antenna," Computers, Materials & Continua, vol. 73, no.1, pp. 917–933, 2022.
[9] R.D. Kulkarni, "Using esemble methods for improving classification of the KDD CUP'99 data set", IOSR Journal of Computer Engineering (IOSR-JCE), vol. 16, pp. 57-61, 2014.
[10] A. A. Abdelhamid and S. R. Alotaibi, "Robust prediction of the bandwidth of metamaterial antenna using deep learning," Computers, Materials & Continua, vol. 72, no.2, pp. 2305–2321, 2022.
[11] Y. Zhang, H. Zhang, J. Cai, B. Yang, "A weighted voting classifier based on differential evolution", Abstract and Applied Analysis, 2014.
[12] T. Saito, "Theoretical Model: Condorcet's Jury Theorem, Part 1", Wolfram Demonstrations Project, http://demonstrations.wolfram.com/TheoreticalModelCondorcetsJuryTheoremPart1, 2017.
[13] F. Leon, C.G. Piuleac, S. Curteanu, I. Poulios, "Instance-based regression with missing data applied to a photocatalytic oxidation process", Central European Journal of Chemistry, vol. 10, no. 4, pp. 1149-1156, 2012.
[14] F. Leon, C. Lisa, S. Curteanu, "Prediction of the liquid crystalline property using different classification methods", Molecular Crystals and Liquid Crystals, vol. 518, pp. 129-148, 2010.
[9] F. Leon, S. Curteanu, C. Lisa, N. Hurduc, "Machine learning methods used to predict the liquid-crystalline behavior of some copolyethers", Molecular Crystals and Liquid Crystals, vol. 469, pp. 1-22, Taylor and Francis Group, USA, 2007, DOI: 10.1080/15421400701431232.
[10] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I.H. Witten, "The WEKA data mining software: An Update", ACM SIGKDD Explorations, vol. 11, no. 1, pp. 10–18, 2009.
[11] R.N. Shepard, "Toward a universal law of generalization for psychological science", Science, vol. 237, pp. 1317–1323, 1987.
[12] K.Q. Weinberger and L.K. Saul, "Distance metric learning for large margin nearest neighbor classification", Journal of Machine Learning Research, vol. 10, pp. 207–244, 2009.
[13] F. Leon, S. Curteanu, "Large margin nearest neighbour regression using different optimization techniques", Journal of Intelligent & Fuzzy Systems, vol. 32, pp. 1321-1332, 2017.
[14] Abdelhamid, A.A.; El-Kenawy, E.-S.M.; Khodadadi, N.; Mirjalili, S.; Khafaga, D.S.; Alharbi, A.H.; Ibrahim, A.; Eid, M.M.; Saber, M. Classification of Monkeypox Images Based on Transfer Learning and the Al-Biruni Earth Radius Optimization Algorithm. Mathematics 2022, 10, 3614.
[15] Eid, M.M.; El-Kenawy, E.-S.M.; Khodadadi, N.; Mirjalili, S.; Khodadadi, E.; Abotaleb, M.; et al., Meta-Heuristic Optimization of LSTM-Based Deep Network for Boosting the Prediction of Monkeypox Cases. Mathematics 2022, 10, 3845.

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022

[16] R. A. Fisher, “The use of multiple measurements in taxonomic problems”, Annual Eugenics, vol. 7, part II, pp. 179-188, 1936.
[17] R. O. Duda, P. E. Hart,” Pattern classification and scene analysis”, John Wiley & Sons, p. 218,1973.
[18] I.W. Evett and E.J. Spiehler, “Rule induction in forensic science”, Central Research Establishment. Home Office Forensic Science Service. Aldermaston, Reading, Berkshire RG7 4PN.
[19] P. W. Frey and D. J. Slate, "Letter recognition using holland-style adaptive classifiers", Machine Learning, vol 6, 1991.
[20] D. Sami Khafaga, A. Ali Alhussan, E. M. El-kenawy, A. Ibrahim, S. H. Abd Elkhalik et al., "Improved prediction of metamaterial antenna bandwidth using adaptive optimization of lstm," Computers, Materials & Continua, vol. 73, no.1, pp. 865–881, 2022.

Doi: https://doi.org/10.54216/JAIM.020101
Received: April 06, 2022 Accepted: October 08, 2022