GA

by MA ISLAM

Improving the Accuracy of Machine Learning Models for Predicting School Attrition among Female Students in Bangladesh by Employing Feature Selection Techniques

Mohammad Nayeem Hasan
Department of Statistics
Shahjalal University of Science and Technology
Sylhet, Bangladesh
nayeem5847@gmail.com

Shomi Khan
Department of Electrical and Electronic Engineering
Shahjalal University of Science and Technology
Sylhet, Bangladesh
nkskl6@gmail.com

Sabrin Sultana
Department of Banking and Insurance
University of Chittagong
Chittagong, Bangladesh
sabrinsultana1060@gmail.com

***Abstract*— In real-world scenarios, such as school attrition, it is important to consider relevant features when applying machine learning algorithms. However, not all features are essential, and including unnecessary or irrelevant ones can diminish the effectiveness of the algorithm. Feature selection (FS) is a crucial task in machine learning to reduce data dimensionality and improve algorithm performance. Evolutionary algorithms (EA) have gained significant attention as a means to solve FS problems. Surprisingly, no previous research has applied these methods specifically to predict factors related to school attrition. Therefore, the objective of this study is to employ various machine learning techniques optimized by a genetic algorithm (GA) and compare their performance. Data for this study were collected from the Multiple Indicator Cluster Surveys, 2019. Among the methods tested, a GA combined with an artificial neural network (ANN) demonstrated the highest accuracy in predicting school attrition. Without the GA optimization, the prediction accuracy was lowest across all the applied methods. However, with GA, the accuracy improved from 76.00% to 76.50% in ANN, from 75.25% to 76.00% in support vector machine (SVM), and from 74.92% to 75.83% in logistic regression (LR). Notably, the LR model achieved the best recall at 95.39%, surpassing the baseline of 50%. Additionally, the LR model yielded the highest F1 score, while all methods exhibited results above 50%. These findings indicate that all models performed well in accurately classifying students as in-school or out-of-school. Thus, tailoring filter measures based on specific FS technique characteristics can significantly enhance efficiency and effectiveness.**

***Keywords*—feature selection, evolutionary algorithms, genetic algorithm, machine learning, classification**

I. Introduction

Girls play a vital role in fostering a nation's progress as they lay the foundation for education. Educated women tend to have higher incomes, are more confident in their decision-making, and raise healthier and better-educated children, surpassing their male counterparts in these aspects [1]. In numerous low-income countries at present, there remains a substantial proportion of females who face school dropouts. As indicated by the graph, over 50% of girls commence their primary education but are unable to successfully complete it [2]. A student's family poverty level exerts a significant influence on their likelihood of experiencing attrition in education. Among the factors that contribute to high attrition rates in secondary schools, school-to-school transitions play a prominent role. The academic and social environment of the school directly impacts institutional attrition rates [3], [4].

The process of selecting a subset of features in feature selection (FS), which aims to reduce data volume and enhance the efficiency of classification algorithms, poses a challenging task [5]. Reducing the amount of data, particularly by selecting a limited number of relevant features, contributes to the development of improved models characterized by higher accuracy. In the domains of machine learning (ML), data mining, and statistics, evolutionary computations (ECs) provide effective solutions for optimizing problems by emulating certain aspects of biological evolution. Evolutionary algorithms (EA) adopt the strategies derived from ECs, benefiting from the inspiration drawn from biological evolution. Notable optimization techniques employed by EA include principal component (PC) analysis, particle swarm (PS) optimization, ant colony (AC) optimization, and genetic algorithm (GA) optimization [6]. The objective of this study is to employ two distinct machine learning methods, namely artificial neural networks and support vector machines, which have been trained and structurally optimized using genetic algorithms (GA). The results obtained from these methods were compared with regression-based approaches, specifically logistic regression models.

II. Literature Review

A considerable body of research has utilized genetic algorithms (GA) to effectively classify models across various domains. For instance, Barrios et al. [7] utilized GA to train a network and developed a classification technique for breast cancer classification. Another study focused on utilizing GA for feature selection in support vector machines (SVM) to classify defective and normal instances [8]. GA has also been applied by Beiko and Charlebois [9] to identify optimal outputs for artificial neural networks (ANN) in the classification of DNA sequences. Similarly, Karzynski et al. [10] employed GA to reduce data for weight classification in ANN for microarray classifications. Additionally, Cho et al. [11] applied GA to enhance the prediction and classification of cDNA microarrays using ANN outputs.

III. Methods

A. Study Design

This study utilized data from a cross-sectional survey, specifically the Multiple Indicator Cluster Surveys conducted in 2019 (source: https://mics.unicef.org/surveys). The response variable of interest is school attrition, categorized as "Yes" or "No." The explanatory variables used in this study are consistent with those included in previously published research [12], [13].

B. Data Preparation

The entire dataset was evenly divided into two subsets: the training data and the validation (testing) data. The validation data played a crucial role in identifying any issues that might arise in the classification models, such as overfitting or premature termination of the training process. To address potential outliers in the data matrix, this study employed the standardization of Z-scores.

IV. FEATURE SELECTION METHODS

Feature selection (FS) holds significant importance not only in large datasets but also in dealing with various challenges. One such challenge arises from frequent feature interactions observed in many cases. Features can exhibit diverse patterns, ranging from general characteristics to complex multi-way interactions [14]. Furthermore, an appropriate feature can effectively serve multiple purposes concurrently when utilized in parallel. However, the primary objective of FS is to identify the key terms or features [15], facilitate the creation of graphical representations in image analysis [16], and determine crucial business indicators [17].

A. Genetic Algorithm

The concept of genetic algorithms (GA) was initially proposed by [18] as a technique for addressing a wide range of optimization problems [19]. This method's design draws inspiration from the observation of natural Darwinian evolutionary processes and the principle of survival of the fittest [20]. GA involves the implementation of several parameters, including mutation probability, crossover probability, and their interrelationships [21]. In order to comprehend the underlying theory of genetic algorithms, it is important to grasp the process of recombination. Recombination involves two parents and can be illustrated by considering a binary string, such as 110100110010101, and another binary string represented as yxyyxyxxyyyxyxxy, where x and y represent 0 and 1. Recombination occurs using two "break-points" as follows:

11010	01100101	101
Yxyyx	yxxyyyxy	xxy

By interchanging the fragments between the two parents, the offspring can be generated as: 11010yxxyyyxy101 and yxyyx01100101xxy.

GA can be effectively utilized for training purposes in logistic regression (LR), artificial neural networks (ANN), and support vector machines (SVM), as well as for determining optimal model structures. Notably, GA does not rely on gradient-based error functions, making it less susceptible to the problem of local minima [22]. The step-by-step process of GA is depicted in Figure 1. Initially, a set of randomly generated codes is created and divided into an initial population. Subsequently, fitness values are assigned to each code by evaluating the fitness function. This is followed by performing a crossover, which involves exchanging features among the selected subset, and introducing mutations that randomly modify selected features. The process then loops back to the second step and continues until the exit criteria are met.

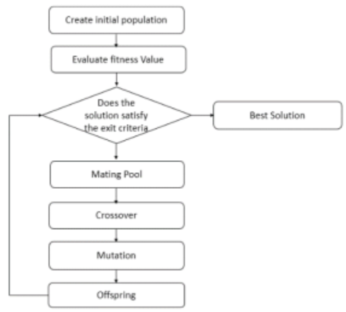

Figure. 1. Steps of genetic algorithm

Initial Population: The size of the initial population plays a crucial role in GA. It is determined based on problem-specific criteria and is typically initialized with randomly generated trial solutions.

Fitness Computation: The computation of fitness is a vital component in the classification process. It involves evaluating and assigning appropriate fitness values to the input samples in our models. This step helps determine the fitness or suitability of each individual in the population.

Crossover: During the crossover phase, the encoded bit string of a specific length is divided and exchanged between parental variables. The exact number of bits before and after a specific point position is swapped, facilitating the generation of new offspring with potentially improved characteristics.

Mutation: Mutations occur with a small probability, often set at values like 0.001. A 100% mutation rate would result in a complete reformulation of the entire dataset, while a 0% mutation rate implies that no part of the dataset will be altered. Mutations introduce small, random changes to the genetic makeup of individuals in the population, contributing to diversity and potentially leading to improved solutions.

Initial Population: The initial population size are the major features in GA. Initialization is known as the size of population which usually selected by the criteria of the problem by initially generated random trial.

B. Application of Genetic Algorithm

1) **Population Initialization:** Create an initial population consisting of randomly generated bit strings with binary values.

2) **Decoding:** Decode the bit strings to determine which input variables are selected.

3) **Classification Model:** Apply classification models during this stage.

4) **Fitness Evaluation:** Evaluate the fitness of each model by assessing its prediction accuracy, which serves as the fitness metric for GA.

5) **Stopping Criterion:** Determine whether the loop should continue or exit based on predefined stopping criteria.

6) **Selection:** Randomly select individuals from the population for crossover, indicating which models will undergo the crossover process.

7) **Crossover:** Apply a crossover operator to produce a linear combination of the selected models.

8) **Mutation:** Introduce a new gene through the same operator, randomly selecting a subset of crossover variables.

9) **Replacement:** Replace the old variables with the two best combinations generated from the mutation step, preparing them for the next iteration.

10) **Loop:** Return to Step 2 and repeat the process.

C. Prediction Models

1) **Artificial Neural Network (ANN):** ANN presents an alternative approach for comprehending and identifying intricate patterns within datasets. It can be viewed as an iterative process that converges towards solutions for various classification problems.

2) **Support Vector Machine (SVM):** SVM is a method specifically designed to handle noisy patterns within large datasets. This technique involves finding a function in a multidimensional space that effectively separates the training data based on known class labels.

3) **Logistic Regression (LR):** LR is utilized to analyze binary response data, where the outcome variable Y is represented by either 1 ("success") or 0 ("failure"). The logistic model for a variable Y is defined as

$$1/(e^{(-Y)}+1) \qquad (1)$$

D. Fitness Evaluation

The fitness values in GA are evaluated based on the accuracy metric, which is calculated using the formula:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Here, TP refers to true positives, FP represents false positives, TN indicates true negatives, and FN corresponds to false negatives. True positives (TP) are instances that are correctly classified as positive. False positives (FP) are instances that are incorrectly classified as positive. True negatives (TN) are instances that are correctly classified as negative, while false negatives (FN) are instances that are incorrectly classified as negative.

Additionally, other performance metrics can be derived from these values. Recall (R) is calculated as TP / (TP + FN), representing the ability of the model to correctly identify positive instances. Precision (P) is computed as TP / (TP + FP), indicating the proportion of correctly classified positive instances out of all instances predicted as positive. F1-score is a harmonic mean of precision and recall, given by (2 * P * R) / (P + R). Therefore, these evaluation metrics provide valuable insights into the performance of the classification models and their effectiveness in accurately classifying instances.

V. RESULTS AND DISCUSSION

A subsample of 4800 girls was generated for the analysis, comprising different age groups. Among them, 15-year-olds accounted for 36.40%, 16-year-olds for 34.10%, and 17-year-olds for 29.50%. Within the group of 15-year-olds, 7.20% were out of school, while 29.20% were attending school. Similarly, the percentages of out-of-school girls for the 16 and 17-year-old groups were 10.60% and 11.30%, respectively.

To obtain the results in this study, the GA was configured as outlined in Table 1. The selection of suitable fitness functions played a crucial role in minimizing classification errors for the classification models. The proximity of the best fitness and mean fitness values to the standard values indicated that the GA had reached the termination condition. The Stall Generation parameter represented the number of generations produced by the GA after the latest improvement in fitness value.

In this study, the GA utilized single-point crossover, a mutation threshold of 0.55, a maximum iteration of 10, a population size of 40, and a number of offspring of 40 in each iteration. Previous research by T. Liu et al. employed a population size of 20 and stopped all algorithms after 200 iterations [23]. The parameters used by the GA for all models can be found in Table I.

TABLE I. PARAMETERS USED IN GA

GA Parameter	Value
Crossover	Single point crossover
Mutation threshold	0.55
Maximum iteration	10
Number of population (Initial)	40
Number of offspring in each iteration	40

The application of the GA feature selector in ANN classification yielded a selection of 6 features, with a mean value of 0.67. Notably, the GA selected 7 features for both SVM and LR models. This finding aligns with the results of a study conducted by Ahmed et al. [24] who used similar parameters and identified the same set of features. It is worth mentioning that the parameters listed in the GA configuration table can still be fine-tuned to improve the results. This characteristic of the GA method provides a higher level of controllability. Furthermore, the mean feature value was consistently higher across all models, as depicted in Table II.

TABLE II. SELECTED FETURE

Feature	ANN	SVM	LR
Girls age (Year)	1	1	1
Marital status	1	1	1
Area	1	0	0
Divisions	1	1	1
Household wealth index	1	0	0
Religion	0	1	1
Household education	0	1	1
Mother alive	1	1	1
Father alive	0	1	1
Mean	0.67	0.78	0.78

Table III presents the training parameters employed in the ANN model. Through extensive experimentation, we determined that the optimal model configuration consisted of four hidden layers with node counts of 6, 9, 9, and 6, respectively. The chosen learning rate for this study was 0.001. In a previous study [25], we had reported the use of a learning rate of 0.01, 10 nodes for the hidden layer, and a 20% dataset allocation for testing.

TABLE III. PARAMETERS USED IN ANN

ANN Parameter	Value
Hidden layer	4 [6,9,9,6]
Solver	adam
alpha	1e-5
activation	relu
output activation	softmax
Learning rate initialize	0.001
Maximum initialize	500

The results clearly indicate that the prediction accuracy of all methods was lower when GA feature selection was not applied. Specifically, without GA, the accuracy rates were 76.00%, 75.25%, and 74.92% for ANN, SVM, and LR, respectively. Conversely, when GA feature selection was incorporated, the accuracy improved to 76.50%, 76.00%, and 75.83% for ANN, SVM, and LR, respectively (see Table 4). These findings align with those of Kim, who reported similar results in their study [26]. Therefore, it is evident that the inclusion of GA feature selection consistently yielded the highest accuracy across all methods [26].

TABLE IV. CLASSSIFICATION ACCURACY WITH AND WITHOUT GA

Classification Method	Without GA	With GA	Training time
ANN	76.00%	76.50%	2.864
SVM	75.25%	76.00%	0.603
LR	74.92%	75.83%	0.043

In the LR model, we achieved the highest recall of 95.39%, which is significantly better than the 50% threshold. Additionally, the LR model exhibited the highest F1 score among all the methods. Importantly, all the models yielded results above 50%, indicating their suitability for our dataset. Thus, it can be concluded that all the models performed well and effectively captured the patterns and characteristics present in our data.

TABLE V. RESULTS OF ATTRITION CLASSIFICATION USING GA ON DIFFERENT METHODS

Methods	Recall (R) (%)	Precession (P) (%)	F1-score (%)
ANN	92.35	77.65	84.36
SVM	93.67	76.59	84.27
LR	95.39	75.72	84.42

In the LR model, we achieved a remarkable recall rate of 95.39%, surpassing the minimum threshold of 50%. Additionally, the LR model exhibited the highest F1 score compared to other methods, indicating its superior performance. Furthermore, all the models demonstrated results above the 50% benchmark, suggesting that they were well-suited for our dataset. Therefore, it can be concluded that all the models effectively captured the underlying patterns and characteristics of our data.

VI. CONCLUSION

In this research, we developed an independent classification method and a GA model to classify school attrition data. The findings of the study indicate that the genetic algorithm yielded favorable results within a limited range of initial parameters. In most cases, the difference in classification accuracy with and without GA was minimal. However, the GA feature selectors consistently produced improved classification accuracy without introducing complexity to the method. An important advantage of this approach is its controllability, as the GA can be fine-tuned to achieve better results by adjusting the fitness functions.

Ultimately, the selection of the best model methodology depends on the application of the model and the likelihood of the required assumptions being true. In conclusion, when choosing which approach to utilize, careful consideration must be given to the assumptions underlying each method.

VII. ETHICS APPROVAL

This research utilized a secondary dataset obtained from the UNICEF website, which is publicly accessible and does not contain any identifying information. As a result, institutional ethics approval was not required for this study (https://mics.unicef.org/surveys).

REFERENCES

[1] M. A. Amadi, E. Role, and L. N. Makewa, "Girl Child Dropout: Experiential Teacher and Student Perceptions," *International Journal of Humanities and Social Science*, vol. 3, no. 5, p. 8, 2013, [Online]. Available: http://www.ijhssnet.com/journals/Vol_3_No_5_March_2013/13.pdf.

[2] R. Sabates, A. Hossain, and K. M. Lewin, *Consortium for Research on Educational Access, Transitions and Equity School Drop Out in Bangladesh: New Insights from Longitudinal Evidence*, no. 49. 2010.

[3] J. W. Alspaugh, "The Effect of Transition Grade to High School, Gender, and Grade Level Upon Dropout Rates," *American Secondary Education*, vol. 29, no. 1, pp. 2–9, 2000, doi: 10.2307/41064411.

[4] P. Goldschmidt and J. Wang, "When can schools affect dropout behavior? A longitudinal multilevel analysis," *American Educational Research Journal*, vol. 36, no. 4, pp. 715–738, 1999, doi: 10.3102/00028312036004715.

[5] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A Survey on Evolutionary Computation Approaches to Feature Selection," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 606–626, Aug. 2016, doi: 10.1109/TEVC.2015.2504420.

[6] B. de la Iglesia, "Evolutionary computation for feature selection in classification problems," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 3, no. 6, pp. 381–407, Nov. 2013, doi: 10.1002/widm.1106.

[7] D. Barrios, A. Carrascal, D. Manrique, and J. Ríos, "Cooperative binary-real coded genetic algorithms for generating and adapting artificial neural networks," *Neural Computing and Applications*, vol. 12, no. 2, pp. 49–60, Nov. 2003, doi: 10.1007/s00521-003-0364-1.

[8] B. Samanta, K. R. Al-Balushi, and S. A. Al-Araimi, "Bearing Fault Detection Using Artificial Neural Networks and Genetic Algorithm," *Eurasip Journal on Applied Signal Processing*, vol. 2004, no. 3, pp. 366–377, Mar. 2004, doi: 10.1155/S1110865704310085.

[9] R. G. Beiko and R. L. Charlebois, "GANN: Genetic algorithm neural networks for the detection of conserved combinations of features in DNA," *BMC Bioinformatics*, vol. 6, no. 1, p. 36, Feb.

2005, doi: 10.1186/1471-2105-6-36.

[10] M. Karzynski, Á. Mateos, J. Herrero, and J. Dopazo, "Using a genetic algorithm and a perceptron for feature selection and supervised class learning in DNA microarray data," *Artificial Intelligence Review*, vol. 20, no. 1–2, pp. 39–51, Oct. 2003, doi: 10.1023/A:1026032530166.

[11] H. S. Cho, T. S. Kim, J. W. Wee, S. M. Jeon, and C. H. Lee, "cDNA microarray data based classification of cancers using neural networks and genetic algorithms," in *2003 Nanotechnology Conference and Trade Show - Nanotech 2003*, 2003, vol. 1, pp. 28–31, Accessed: Aug. 07, 2020. [Online]. Available: https://briefs.techconnect.org/papers/cdna-microarray-data-based-classification-of-cancers-using-neural-networks-and-genetic-algorithms/.

[12] M. N. Hasan, "A Comparison of Logistic Regression and Linear Discriminant Analysis in Predicting of Female Students Attrition from School in Bangladesh," Dec. 2019, doi: 10.1109/EICT48899.2019.9068776.

[13] M. N. Hasan, "Factors Associated with Attrition of Girls Students from School in Bangladesh," *Journal of Scientific Research*, vol. 12, no. 1, pp. 29–38, Jan. 2020, doi: 10.3329/jsr.v12i1.41579.

[14] I. Guyon and A. M. De, "An Introduction to Variable and Feature Selection André Elisseeff," 2003.

[15] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text feature selection using ant colony optimization," *Expert Systems with Applications*, vol. 36, no. 3 PART 2, pp. 6843–6853, Apr. 2009, doi: 10.1016/j.eswa.2008.08.022.

[16] A. Ghosh, A. Datta, and S. Ghosh, "Self-adaptive differential evolution for feature selection in hyperspectral image data," *Applied Soft Computing Journal*, vol. 13, no. 4, pp. 1969–1977, Apr. 2013, doi: 10.1016/j.asoc.2012.11.042.

[17] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 4, pp. 491–502, Apr. 2005, doi: 10.1109/TKDE.2005.66.

[18] J. H. Holland, "Adaptation in Natural and Artificial Systems | The MIT Press," *The University of Michigan Press*, 1975. https://mitpress.mit.edu/books/adaptation-natural-and-artificial-systems (accessed Jul. 31, 2020).

[19] H. Chiroma *et al.*, "Neural networks optimization through genetic algorithm searches: A review," *Applied Mathematics and Information Sciences*, vol. 11, no. 6. pp. 1543–1564, 2017, doi: 10.18576/amis/110602.

[20] M. Hamdan, "A heterogeneous framework for the global parallelisation of genetic algorithms," *International Arab Journal of Information Technology*, vol. 5, no. 2, pp. 192–199, 2008, Accessed: Jul. 31, 2020. [Online]. Available: https://www.researchgate.net/publication/220413569_A_Heterogeneous_Framework_for_the_Global_Parallelisation_of_Genetic_Algorithms.

[21] C. T. Capraro, I. Bradaric, G. T. Capraro, and T. K. Lue, "Using genetic algorithms for radar waveform selection," 2008, doi: 10.1109/RADAR.2008.4720947.

[22] "BP Neural Network Algorithm Optimized by Genetic Algorithm and Its Simulation," *International Journal of Computer Science Issues*, vol. 10, no. 1, pp. 516–519, 2013, Accessed: Jul. 31, 2020. [Online]. Available: https://www.researchgate.net/publication/303102486_BP_Neural_Network_Algorithm_Optimized_by_Genetic_Algorithm_and_Its_Simulation.

[23] T. Liu, H. Zhang, H. Zhang, and A. Zhou, "Information Fusion in Offspring Generation: A Case Study in Gene Expression Programming," *IEEE Access*, vol. 8, pp. 74782–74792, 2020, doi: 10.1109/ACCESS.2020.2988587.

[24] F. Ahmad, N. A. Mat-Isa, Z. Hussain, R. Boudville, and M. K. Osman, "Genetic Algorithm - Artificial Neural Network (GA-ANN) hybrid intelligence for cancer diagnosis," in *Proceedings - 2nd International Conference on Computational Intelligence, Communication Systems and Networks, CICSyN 2010*, 2010, pp. 78–83, doi: 10.1109/CICSyN.2010.46.

[25] Y. T. Chang, J. Lin, J. S. Shieh, and M. F. Abbod, "Optimization the initial weights of artificial neural networks via genetic algorithm applied to hip bone fracture prediction," *Advances in Fuzzy Systems*, 2012, doi: 10.1155/2012/951247.

[26] K. Kim, "Artificial neural networks with evolutionary instance selection for financial forecasting," *Expert Systems with Applications*, vol. 30, no. 3, pp. 519–526, Apr. 2006, doi: 10.1016/j.eswa.2005.10.007.

GA

ORIGINALITY REPORT

32%	27%	29%	18%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Mohammad Nayeem Hasan. "A Comparison of Logistic Regression and Linear Discriminant Analysis in Predicting of Female Students Attrition from School in Bangladesh", 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019 Publication	2%
2	www.hindawi.com Internet Source	1%
3	eprints.bournemouth.ac.uk Internet Source	1%
4	S. Sarath, P.Sam Paul, G. Lawrance, D.S.Shylu Sam. "Optimizing MRF foam Damper Parameters using Artificial Neural Networks and Genetic Algorithms for Improved Damping Force Performance", 2023 4th International Conference on Signal Processing and Communication (ICSPC), 2023 Publication	1%

Submitted to Stockholms universitet

5	Student Paper	1%
6	ijeecs.iaescore.com Internet Source	1%
7	www.ijeast.com Internet Source	1%
8	Davies Segera, Mwangi Mbuthia, Abraham Nyete. "Metaheuristics for optimal feature selection in high-dimensional datasets", Elsevier BV, 2023 Publication	1%
9	Mohammadamin Esmaeili, Mohammad Reza Moradi, Hamid Reza Afshoun. "A new empirical model and neural network-based approach for evaluation of isobaric heat capacity of natural gas", Journal of Natural Gas Science and Engineering, 2022 Publication	1%
10	Shomi Khan, M. Elieas Ali, Sourav Das, Md Mohsinur Rahman. "Real Time Hand Gesture Recognition by Skin Color Detection for American Sign Language", 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019 Publication	1%
	www.slideshare.net	

11	Internet Source	1 %
12	Whitley, D.. "Genetic algorithms and neural networks: optimizing connections and connectivity", Parallel Computing, 199008 Publication	1 %
13	www.thefreelibrary.com Internet Source	1 %
14	theses.hal.science Internet Source	1 %
15	Neema Mduma, Dina Machuve. "Machine Learning Model for Predicting Student Dropout: A Case of Tanzania, Kenya and Uganda", 2021 IEEE AFRICON, 2021 Publication	1 %
16	ijcsi.org Internet Source	1 %
17	researchportal.hw.ac.uk Internet Source	1 %
18	Hu Zhang, Hengzhe Zhang, Aimin Zhou. "A Multi-metric Selection Strategy for Evolutionary Symbolic Regression", 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020 Publication	1 %
19	ijece.iaescore.com	

Internet Source

1%

20	Dong-Ling Tong. "Hybridising Genetic Algorithm-Neural Network (GANN) in marker genes detection", 2009 International Conference on Machine Learning and Cybernetics, 2009 Publication	1%
21	www.researchgate.net Internet Source	1%
22	core.ac.uk Internet Source	1%
23	researchpub.org Internet Source	1%
24	www.ru.ac.bd Internet Source	1%
25	Prasanjit Chakraborty, Sukanta Nama, Apu Kumar Saha. "A hybrid slime mould algorithm for global optimization", Multimedia Tools and Applications, 2022 Publication	1%
26	Submitted to College of Southern Nevada, West Charleston Campus Student Paper	1%
27	www.sersc.org Internet Source	1%

28	kar.kent.ac.uk Internet Source	1 %
29	repositorio.unal.edu.co Internet Source	1 %
30	m.scirp.org Internet Source	<1 %
31	ndl.ethernet.edu.et Internet Source	<1 %
32	Montri Inthachot, Veera Boonjing, Sarun Intakosum. "Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend", Computational Intelligence and Neuroscience, 2016 Publication	<1 %
33	research.library.mun.ca Internet Source	<1 %
34	Submitted to Swinburne University of Technology Student Paper	<1 %
35	Ojak Abdul Rozak, Mohd Zamri Ibrahim, Muhamad Zalani Daud, Syaiful Bakhri, Rifqi Muwaffiq. "Impact of cell temperature on the performance of a rooftop photovoltaic system of 2.56 kWp at Universitas Pamulang",	<1 %

Indonesian Journal of Electrical Engineering and Computer Science, 2023
Publication

---

36 Wanru Li, Jeremy M. Manheim, Yue Fu, Tiina Laaksonen, Gozdem Kilaz, Hilkka I. Kenttämaa. "Comparison of APCI orbitrap MS and GCxGC/EI TOF MS for the hydrocarbon analysis of heavy base oils", Fuel, 2023
Publication
<1 %

---

37 hal.archives-ouvertes.fr
Internet Source
<1 %

---

38 link.springer.com
Internet Source
<1 %

---

39 www.ijcsi.org
Internet Source
<1 %

---

40 Md. Monirul Kabir, Md. Shahjahan, Kazuyuki Murase. "A new hybrid ant colony optimization algorithm for feature selection", Expert Systems with Applications, 2012
Publication
<1 %

---

41 fenix.tecnico.ulisboa.pt
Internet Source
<1 %

---

42 www.scirus.com
Internet Source
<1 %

---

43 Aya Hesham, Nora El-Rashidy, Amira Rezk, Noha A. Hikal. "Towards an Accurate Breast
<1 %

Cancer Classification Model based on Ensemble Learning", International Journal of Advanced Computer Science and Applications, 2022
Publication

44	Daniel Bustos Coral, Maristela Oliveira Santos, Claudio Fabiano Motta Toledo. "Clustering-Based Crossover in an Evolutionary Algorithm for the Vehicle Routing Problem with Time Windows", 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), 2017 Publication	<1 %
45	Ghosh, Ashish, Aloke Datta, and Susmita Ghosh. "Self-adaptive differential evolution for feature selection in hyperspectral image data", Applied Soft Computing, 2013. Publication	<1 %
46	Sarala Pappula, Teja Nadendla, Nagendra Babu Lomadugu, Sai Revanth Nalla. "Detection and Classification of Pneumonia Using Deep Learning by the Dense Net-121 Model", 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), 2023 Publication	<1 %
47	Suman Yadav, Gourav Kumar Suman, Ram Krishna Mehta. "Study of Electromagnetic	<1 %

Forces on Windings of High Voltage Transformer during Short Circuit Fault", 2020 3rd International Conference on Energy, Power and Environment: Towards Clean Energy Technologies, 2021
Publication

48	Syaharuddin, Fatmawati, Herry Suprajitno. "Experimental Analysis of Training Parameters Combination of ANN Backpropagation for Climate Classification", Mathematical Modelling of Engineering Problems, 2022 Publication	<1 %
49	docshare.tips Internet Source	<1 %
50	ijeee.iust.ac.ir Internet Source	<1 %
51	turcomat.org Internet Source	<1 %
52	uhra.herts.ac.uk Internet Source	<1 %
53	upcommons.upc.edu Internet Source	<1 %
54	www.naturalspublishing.com Internet Source	<1 %
55	hdl.handle.net Internet Source	<1 %

56	B Samanta, Khamis R Al-Balushi, Saeed A Al-Araimi. "Bearing Fault Detection Using Artificial Neural Networks and Genetic Algorithm", EURASIP Journal on Advances in Signal Processing, 2004 Publication	<1 %
57	Dalila Amara, Latifa Rabai. "Comparison of Feature Selection via Semi supervised denoising autoencoder and traditional approaches For Software Fault-prone Classification", Research Square Platform LLC, 2023 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

GA

PAGE 1

S/V This subject and verb may not agree. Proofread the sentence to make sure the subject agrees with the verb.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

PAGE 2

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

S/V This subject and verb may not agree. Proofread the sentence to make sure the subject agrees with the verb.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

PAGE 3

PAGE 4

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

PAGE 5