

Hybrid optimized multi-objective honey badger algorithm and NSGA-II for feature selection problems

Anusha Papasani, Nagaraju Devarakonda

School of Computer Science and Engineering, VIT-AP University, Amaravati, India

Article Info

Article history:

Received Nov 23, 2023

Revised May 27, 2024

Accepted Jun 5, 2024

Keywords:

Honey Badger algorithm
Multi-objective optimization
NSGA-II
Wrapper feature selection

ABSTRACT

One of the most important aspects of classification is choosing features in such a way as to get rid of redundant or irrelevant elements in the dataset. For the most part, multi-objective feature selection strategies have been offered by a number of scholars as a strategy for this aim. On the other hand, these techniques frequently fail to simultaneously improve classification accuracy while removing redundant feature combinations. This article presents a wrapper-based feature selection strategy that strikes a compromise between classification accuracy and redundancy reduction by combining features of the multi objective (MO) based honey badger algorithm (MO-HBA) and non-dominated sorting genetic algorithm-II (NSGA-II). The technique was developed as part of this investigation. Increasing the accuracy of the classification while simultaneously reducing the number of redundant characteristics is one of the optimizations aims of this approach. The MO-HBA shows excellent performance in exploration and exploitation. A Kernel version of the extreme learning machine (KELM) is used for the process of selecting the features to use. In order to evaluate how well this method of feature selection performs, eighteen benchmark datasets are utilized, and the results are compared to four established methods of multi-objective feature selection based on different metrics.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Anusha Papasani
School of Computer Science and Engineering, VIT-AP University
Near Vijayawada - 522237, Andhra Pradesh, India
Email: anoosha.papasani@gmail.com

1. INTRODUCTION

Classification is crucial to data mining and machine learning. While obtaining and constructing datasets with the most relevant features is difficult, the features typically contain redundancies and unnecessary information for classifiers [1], [2]. Optimal feature selection is essential for dataset simplification. This powerful dimensionality reduction method improves accuracy, interpretability, generalization, and overfitting [3]. In general, feature selection algorithms are filters or wrappers. Filter approaches use rough theory, distance, and information theory to choose essential features from intrinsic data [4]. In contrast, wrapper techniques use learning algorithms and various search tactics to discover important traits [5]. Wrappers perform better than filters in classification [6], but they take longer for large datasets [7]. The multi-objective approach of wrapper feature selection minimizes feature subset size and maximizes classification accuracy [8]. Meta-heuristic optimization approaches are popular in real-world applications due to their powerful search capabilities across a vast solution space [9].

Meta-heuristic optimization algorithms are crucial to wrapper algorithms. Many research uses grey wolf optimization (GWO), particle swarm optimization (PSO), genetic algorithm (GA), bat algorithm (BA), and whale optimization algorithm (WOA) [10]. These evolutionary algorithms also have multi-objective and

single-objective techniques, with the former performing better [11], [12]. Single-objective optimization techniques like PSO and GA, HBA [13] may perform poorly in defined locations despite their worldwide search prowess. SPEA2, NSGA-II, and PESA excel at multi-objective functions, lowering error rates and deleting duplicate features [14]-[16]. Given these factors, multi-objective optimization is preferred. This work introduces MOHBNSGA2, a feature selection-focused multi-objective optimization technique combining HBA and NSGA2. MOHBNSGA2's initialized by population and parameters. The synthetic minority over-sampling technique (SMOTE) is used in preprocessing to improve classifier performance on imbalanced datasets. The chosen classifier, kernel extreme learning machine (KELM), is effective and fast. KELM excels at classifying huge datasets with numerous labels. The proposed method is compared to MOBGA-AOS, MO-PSO, BCNSG3, and MWOA feature selection algorithms.

Summary of MOHBNSGA2-based feature selection's main contributions:

- Integrating NSGA2 (MOHBNSGA2) into the multi-objective feature selection algorithm improves HBA's performance.
- KELM categorization performance improved with selected features.
- The proposed feature selection algorithm's accuracy, precision, recall, micro F1-score, selected feature count, training duration must be shown to work.

2. RELATED WORK

As number of feature selection algorithms exist, the vast majority of them view feature selection as a problem with just one objective. Few tackle the problem of selecting features with several objectives. The ant colony optimisation algorithm with Levy Flight was introduced by Liu and Cao [17]. A convolutional neural network (CNN) utilising SMOTE and the Gaussian Mixture Model was suggested by Liu and Cao [17]. There was also a recommendation for a regularised ELM-based quick conformal prediction system. An online feature selection method based on multi-objective PSO was proposed by Paul *et al.* [18] for use in multi-label classification. Using mutual information and an entropy-based gain ratio, four feature selection strategies based on multi-objective filters were suggested by Usman *et al.* [19]. These strategies made use of the multi-objective binary cuckoo optimisation algorithm and NSGA III. Binary differential evolution was the basis for the self-learning approach presented by Zhang *et al.* [20]. A multi-objective feature selection strategy based on wrappers was proposed by Kozodoi *et al.* [21] utilising NSGA-II. A wrapper-based multi-objective feature selection method using NSGA-II and breeding operators was presented by Gonzalez *et al.* [22]. The salp swarm technique and the multi-objective spotted hyena optimizer were coupled for feature selection by Sharma and Rani [23]. One method for selecting features that should be considered in a multi-objective context was suggested by Kiziloz *et al.* [24]. A multi-objective technique for feature ranking was proposed by Amoozegar and Minaei-Bidgoli [25] using PSO. A multi-objective artificial bee colony was used for feature selection by Hancer *et al.* [26] in conjunction with non-dominated sorting genetic operators. A multi-objective approach to gene selection and microarray classification was presented by Dashtban *et al.* [27]. As a strategy for selecting genes, Lai introduced a hybrid filter-wrapper-based multi-objective optimisation approach [28]. One such method that was utilised for selecting features was the elitism-based multi-objective differential evolution approach [29]. Using the binary genetic algorithm, Gao *et al.* [30] introduced a method for selecting features that take into account several objectives. In order to choose text features, Labani *et al.* presented a relative discriminative criterion (MORDC) method with a multi-objective function. An HCSO-based multi-objective optimisation method for feature selection was described by Xiao-Zhi *et al.* For healthcare datasets, Rostami *et al.* [31] suggested an improved multi-objective PSO-based feature selection method. by fusing the fruitfly optimisation algorithm with the harris hawk's optimisation, abdollahzadeh, and gharehchopogh demonstrated a multi-objective approach [32].

3. PRELIMINARIES

3.1. Honey badger algorithm

The honey badger optimization (HBO) algorithm [13] is inspired by the behavior of honey badgers, known for their intelligence and adaptability in finding food sources. This algorithm is a nature-inspired optimization technique used for solving optimization problems. Initialize a population of potential solutions, called honey badgers, randomly within the search space. Each honey badger represents a candidate solution to the optimization problem. Evaluate the objective function for each honey badger to determine its fitness value. The objective function is the function that is being optimized, and its value shows the quality of the solution represented by the honey badger. Based on their fitness values, honey badgers explore the search space to find better solutions. Exploration is typically achieved through various mechanisms such as random movement, local search, or guided search strategies. Update the positions of honey badgers based on their

exploration results and the optimization strategy employed. This step aims to improve the quality of solutions over iterations.

Where x_{prey} stands in for the overall best position. The honey bardger’s capacity for food exploration is symbolized by the symbol ‘ β ’ and ‘ d_i ’ is a representation of the separation between the prey and the honey badger. The variables r_3, r_4, r_5 represents the random numbers between 0 and 1. Flag ‘ F ’ that works as represents the search direction with (1).

$$F = \begin{cases} 1 & \text{if } r_6 \leq 0.5 \\ -1 & \text{else} \end{cases} \tag{1}$$

r_6 represents a random number between 0 and 1. To replicate the journey to the beehive, we can employ (2) to simulate the collaborative behavior of the honey guide bird and the honey badger.

$$x_{new} = x_{prey} + F \times r_7 \times \alpha \times d_i \tag{2}$$

The random number between 0 and 1 is denoted by ‘ r_7 ’. The updated position of the honey badger is given by x_{new} , while x_{prey} represents the location of the prey. The search behavior is affected by the time-varying parameter (α), which is updated during this stage.

3.2. NSGA-II

NSGA-II stands out as a renowned multi-objective optimization technique applicable to feature selection problem-solving. The algorithm begins by identifying the most suitable solutions from the initial population. “Child solutions” are created through the crossing of “parent solutions” selected from this pool. Mutation introduces unpredictable changes to some solution components. In the event that the termination requirements are not met, the best solutions are found using a non-dominated sorting approach. By measuring and optimizing the crowding distance between solutions, this strategy keeps diversity. As a result, we can identify the sample size for the next assessment step. In the end, the population converges on the set of best, non-dominated solutions known as the Pareto frontier as a result of this sorting mechanism. To adapt NSGA-II for feature selection, individuals must be represented as various feature combinations. Additionally, the number of features should be included as a secondary objective alongside the primary objective, typically related to profitability.

3.3. KELM

The “KELM classifier” is a variant of the extreme learning machine (ELM) algorithm that incorporates kernel methods for nonlinear mapping of input data. Unlike traditional ELM, which uses random feature mappings for dimensionality expansion, these variant employs kernel functions to implicitly transform the input space into a higher-dimensional feature space where the data might be more separable. By utilizing kernel functions, such as radial basis function (RBF) kernels or polynomial kernels, the classifier can effectively handle nonlinear relationships between input features without explicitly defining the feature mappings. This allows the model to capture complex patterns in the data and improve classification performance, particularly when dealing with nonlinearly separable datasets. The KELM can be expressed using (3) according to ridge regression theory.

$$\hat{\beta} = \left(\frac{I}{\lambda} + H^T H \right)^{-1} H^T Y \tag{3}$$

Where $Y = H \beta$ and λ represents the regularization coefficient.

4. PROPOSED FEATURE SEKECTION ALGORITHM

Logistic mapping initialization, NSGA-II, and MO-HBA are all parts of the suggested hybrid MOO algorithm. By choosing evolutionary operators on the fly, our hybrid method strikes a good balance between local and global search. Figure 1 depicts the flowchart of the proposed feature selection algorithm.

4.1. Initialization

The MO-HBA approach necessitates the initialization of a population of particles. The NSGA-II algorithm requires the initialization of a population of individuals to achieve the same objective. The MO-HBA algorithm requires articles, while the NSGA-II algorithm requires a population of individuals.

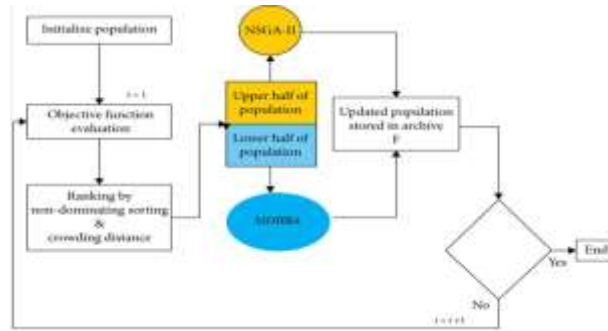


Figure 1. Flowchart of proposed algorithm

4.2. Fitness evaluation

It is necessary to ascertain whether or not the solutions that were generated by the MO-HBA stage are appropriate. To achieve this specific goal, it is important to compute the values of the goal functions for each solution. This evaluation ensures the suitability of the generated solutions.

4.3. Pareto fronts

To the solutions that were produced from the MO-HBA phase, apply the stages of NSGA-II that involve the calculation of crowding distance and non-dominated sorting distance. This will generate a Pareto front. Additionally, crowding distances will be calculated.

4.4. Fitness function

Finding a combination of features that optimizes solution size and classification accuracy is the ultimate aim of the multi-objective feature selection approach. As an initial fitness function, we don't aim to maximize classification accuracy but rather to minimize classification error. The solution size is taken into account by the second fitness function. The solutions were rated using a classifier built from k-nearest neighbors (k-NN). The k-NN is tested by comparing multiple samples. In order to determine the fitness functions, one can use (4), (5).

$$\min(f_1) = \left(\frac{1}{n} \sum_{l=1}^n \frac{N_{Error}}{N_{All}} \right) \times 100\% \quad (4)$$

$$\min(f_2) = \sum_{i=1}^D x_i \quad (5)$$

4.5. Selection of evolutionary operators

Two distinct operators for crossovers (SBX and two-point crossover) and mutations (non-uniform and polynomial mutation) have been the subject of a dynamic selection procedure for evolutionary operators. Non-uniform mutation outperforms polynomial mutation when it comes to local exploitations, and SBX crossover is renowned for its better local search ability hence, linking SBX crossover with non-uniform mutation improves local search ability, while pairing two-point crossover with polynomial mutation improves global search ability. A collection of polynomial mutations and two-point crossovers is chosen to start widening the search space. A non-uniform mutation set and SBX crossover are preferred to promote convergence as the iteration advances, though. In (6) and (7) refers as evolutionary operators.

$$P_S = \frac{T}{T_{max}} \quad (6)$$

$$P_T = 1 - \frac{T}{T_{max}} \quad (7)$$

4.6. Crossover and mutation

In order to generate offspring solutions from the individuals that have been selected, genetic operations should be applied. These operations include crossover and mutation. These processes are essential for creating new solutions.

4.7. Offspring evaluation and selection

Carry out an assessment of the viability of the recently developed offspring solutions. Select individuals for the next generation from the combined set of solutions, considering both parents and offspring. Solutions are chosen based on their non-dominated rank and crowding distance.

4.8. Termination criterion

Check if a termination criterion is met. This could be a maximum number of generations, a specific level of convergence, or any other stopping criterion defined by the user. The pseudo code for proposed feature selection algorithm is detailed in Algorithm 1.

Algorithm 1. Proposed feature selection algorithm-pseudo code

```

Initialization:
Initialize MO-HBA particle and NSGA-II individual populations, Parameters
MO-HBA Phase: Assessment fitness for each solution
                Find pareto fronts
NSGA-II: Non-dominated sorting and crowding distance calculation procedures to MO-HBA
solutions.
Create a Pareto front and
Find out crowding distances.
Selection:
//Select solutions from MO-HBA and NSGA-II solutions based on non-dominated rank and
crowding distance //
Perform Mutation, crossover:
//Create children from selected individuals using crossover and mutation//
Evaluation of Offspring:
Assess the new offspring solutions' fitness.
Selection for Survival:
//Choose parents and children from the integrated =solutions for the next generation//
Solutions are chosen by crowding distance and non-dominated rank.
Reiterate: terminate when all stop conditions are met
Find the error rate for the feature subset and store the results
The population is ranked using non-dominated sorting.
Please return the features that were selected.
    
```

5. EXPERIMENTS

This section details the experimental setup, including the datasets, parameters, and performance metrics [33]. Credit, WDBC, Glass, Ionosphere, Tic-Tac-Toe, Dermatology, Audiology, Lymphography, Sonar, Heart, Spect, Parkinson, Zoo, Indian Pima, Scene, Kc1, Wine, and Waveform are among the datasets employed to assess the efficacy of the suggested method. Table 1 outlines the parameters.

Table 1. Parameters of different algorithms

Method	Parameter
Proposed MOHBNSGA2	No of Honey Badgers=50 β – value =6, C-value=2
MO based WOA (MWOA)	Crossover value= 0.6, Mutation value = 0.05
BCNSG3 –optimization technique	B-value = 0.6, Crossover value = 0.5
MO based BGA-AOS(MOBGA-AOS)	Crossover value= 0.69, Mutation value= 0.05, M-value=2, Q-value=5, LP -Value=5
MO based PSO (MOPSO)	w_{min} = 0.42 w_{max} = 0.93

6. RESULTS

The proposed feature selection strategy demonstrates superior accuracy even with smaller feature sets. Achieving up to 98% accuracy on datasets like Ionosphere, Parkinson, Wine, and Zoo, and over 95% accuracy on WDBC, Dermatology, Sonar, and Audiology datasets. Additionally, datasets such as Lymphographic, Heart, and Scene still maintain accuracy above 90%. Computation times for MWOA, BCNSG3, MOBGA-AOS, and MO-PSO are within 350.58 to 365.59 seconds, showcasing only a slight difference compared to existing approaches.

Several multi-objective feature selection algorithms, including MWOA, BCNSG3, MOBGA-AOS, and MO-PSO, were utilized to evaluate the effectiveness of the MOHBNSGA2 technique. Comparisons across eighteen datasets indicate that the proposed method exhibits lower standard deviation and mean values compared to these other algorithms. The performance of MOHBNSGA2 was evaluated against these known algorithms in terms of accuracy, precision, micro F1-score, recall, and the number of selected features. Tables 2-4 present the performance evaluation results. Across all datasets, MOHBNSGA2 consistently

outperformed the other algorithms, achieving higher values for accuracy, precision, micro F1-score, recall, and specificity, regardless of the data type. Additionally, MOHBNSGA2 selected a smaller number of features, leading to significantly reduced classification error. However, it should be noted that the training time required for MOHBNSGA2 is significantly longer than that of previous methods. Figure 2 shows pareto front analysis of the proposed approach in terms of feature size and error rate values.

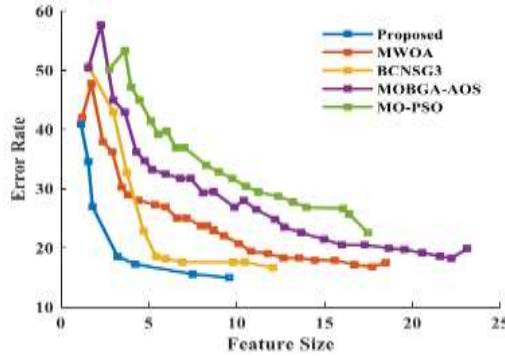


Figure 2. Pareto front analysis of the proposed approach in terms of feature size and error rate

Table 2. The performance comparison was done using MOHBNSGA2 for all the datasets

S.No	Dataset	MOHBNSGA2					
		Accuracy	Precision	Micro F1-score	Recall	Time	Number of selected features
1	WDBC	0.96	0.96	0.95	0.94	9.22870	5
2	Zoo	0.98	0.98	0.98	0.98	1.52199	6
3	Lymphography	0.94	0.94	0.94	0.92	1.97090	5
4	Ionosphere	0.98	0.98	0.98	0.98	3.8346	8
5	Credit	0.85	0.86	0.86	0.85	27.74801	4
6	Heart	0.92	0.92	0.92	0.92	4.26050	8
7	Dermatology	0.97	0.97	0.97	0.99	5.64286	19
8	Sonar	0.96	0.96	0.95	0.95	1.8670	24
9	Spect	0.81	0.82	0.80	0.79	1.71554	9
10	Parkinson	0.98	0.98	0.98	0.98	2.49066	7
11	Indian Pima	0.81	0.82	0.81	0.77	9.00353	4
12	Scene	0.90	0.90	0.89	0.90	67.68548	81
13	Wine	0.98	0.98	0.98	0.98	0.98955	5
14	Tic tac toe	0.86	0.86	0.86	0.86	15.82712	9
15	Glass	0.80	0.82	0.80	0.95	2.46151	4
16	Kc1	0.87	0.87	0.87	0.87	3438222	11
17	Audiology	0.95	0.95	0.95	0.95	1.999073	8
18	Waveform	0.77	0.79	0.79	0.77	375.99	6

Table 3. The performance comparison was done using BCNSG3, MWOA for all the datasets

S.No	Dataset	BCNSG3						MWOA					
		A	P	F1	R	T	N.S.F	A	P	F1	R	T	N.S.F
1	WDBC	0.737	0.74	0.72	0.68	8.28492	7	0.82	0.83	0.81	0.78	5.91051	9
2	Zoo	0.85	0.88	0.85	0.85	1.35818	8	0.76	0.80	0.76	0.75	1.37703	12
3	Lymphography	0.867	0.87	0.86	0.85	1.71169	8	0.89	0.90	0.89	0.87	1.19205	12
4	Ionosphere	0.94	0.94	0.94	0.94	3.60555	7	0.93	0.93	0.93	0.93	3.15333	11
5	Credit	0.85	0.86	0.85	0.86	16.37239	5	0.85	0.85	0.85	0.85	6.80941	10
6	Heart	0.91	0.91	0.91	0.91	4.27747	6	0.89	0.90	0.89	0.89	2.64523	10
7	Dermatology	0.96	0.96	0.96	0.99	5.49611	16	0.93	0.94	0.93	0.99	3.36268	28
8	Sonar	0.94	0.95	0.94	0.94	1.352246	29	0.91	0.91	0.91	0.91	1.42234	39
9	Spect	0.78	0.82	0.78	0.82	1.62678	8	0.78	0.78	0.78	0.76	1.78911	15
10	Parkinson	0.93	0.93	0.93	0.93	1.71446	10	0.97	0.97	0.97	0.97	2.40658	7
11	Indian Pima	0.80	0.80	0.80	0.77	8.70806	6	0.78	0.78	0.76	0.73	5.40530	5
12	Scene	0.89	0.89	0.89	0.89	65.366	152	0.87	0.88	0.87	0.88	62.9994	267
13	Wine	0.96	0.96	0.19	0.96	0.79036	7	0.87	0.90	0.36	0.86	0.47830	10
14	Tic tac toe	0.70	0.73	0.70	0.67	2.67566	4	0.70	0.74	0.70	0.65	2.21453	4
15	Glass	0.78	0.83	0.78	0.96	1.17128	2	0.75	0.79	0.75	0.95	2.37989	3
16	Kc1	0.78	0.79	0.78	0.96	38.2133	8	0.83	0.85	0.83	0.83	36.8245	20
17	Audiology	0.89	0.90	0.89	0.90	1.65898	39	0.83	0.86	0.83	0.85	1.64329	21
18	Waveform	0.66	0.79	0.70	0.66	328.82	20	0.72	0.75	0.72	0.716	350.58	23

Table 4. MOBGA-AOS and MO-PSO performance comparison on all datasets

S.No	Dataset	MOBGA-AOS					MO-PSO						
		A	P	F1	R	T	N.S.F	A	P	F1	R	T	N.S.F
1	WDBC	0.912	0.91	0.91	0.90	4.39573	16	0.877	0.88	0.87	0.85	4.70819	12
2	Zoo	0.84	0.87	0.84	0.84	0.90129	9	0.96	0.97	0.96	0.96	0.90392	8
3	Lymphography	0.92	0.92	0.92	0.90	0.95735	9	0.94	0.94	0.94	0.93	0.95456	7
4	Ionosphere	0.94	0.94	0.94	0.94	2.32057	15	0.90	0.91	0.90	0.90	2.09531	9
5	Credit	0.845	0.85	0.85	0.85	27.37628	6	0.84	0.84	0.84	0.84	10.8155	9
6	Heart	0.894	0.89	0.89	0.89	1.0655	7	0.86	0.86	0.86	0.86	1.12316	7
7	Dermatology	0.93	0.94	0.93	0.99	2.70390	22	0.96	0.96	0.96	0.96	2.55456	21
8	Sonar	0.906	0.91	0.91	0.91	0.87902	26	0.94	0.95	0.94	0.94	0.92037	29
9	Spect	0.78	0.78	0.78	0.76	1.02381	9	0.71	0.71	0.71	0.70	0.99448	13
10	Parkinson	0.97	0.97	0.97	0.97	1.36654	13	0.98	0.98	0.98	0.98	1.31301	6
11	Indian Pima	0.779	0.77	0.77	0.74	5.15370	4	0.80	0.81	0.80	0.76	4.69896	3
12	Scene	0.89	0.89	0.89	0.89	71.5721	157	0.89	0.90	0.89	0.89	68.6277	161
13	Wine	0.93	0.93	0.27	0.93	0.51599	3	0.14	0.14	0.14	0.14	0.50055	4
14	Tic tac toe	0.77	0.77	0.77	0.77	1.43327	5	0.83	0.83	0.83	0.83	5.76226	6
15	Glass	0.78	0.83	0.77	0.96	0.64164	3	0.76	0.79	0.76	0.93	0.68031	4
16	Kc1	0.82	0.83	0.82	0.82	35.6545	9	0.83	0.84	0.83	0.83	39.8770	12
17	Audiology	0.94	0.94	0.94	0.94	0.83690	25	0.893	0.90	0.89	0.90	1.01391	37
18	Waveform	0.68	0.72	0.68	0.68	337.23	19	0.72	0.75	0.72	0.72	365.59	20

*Note: A=accuracy, P=precision, F1: F1-score, R=recall, T=computational time, N.S.F= no of selected features

7. CONCLUSION

In order to minimize classification error while selecting a lower number of features, this research offers a novel multi-objective wrapper feature selection technique. Named MOHBNSGA2, the suggested method integrates two existing algorithms: NSGAL and MOHBA. We compare the method's performance to four well-established multi-objective feature selection algorithms and test it on eighteen benchmark datasets to see how well it works. We use the KELM to assess the features that were chosen. With fewer characteristics picked and reduced classification error, the suggested method achieves better performance than other methods in terms of performance measures. The suggested approach has problems with premature convergence and a somewhat longer runtime than competing methods. Improvements in diversity maintenance and convergence speed will be the primary goals of future efforts.





REFERENCES

- [1] P. Hu, J. S. Pan, and S. C. Chu, "Improved binary grey wolf optimizer and its application for feature selection," *Knowledge-Based Systems*, vol. 195, p. 105746, May 2020, doi: 10.1016/j.knosys.2020.105746.
- [2] Y. Xue, H. Zhu, J. Liang, and A. Slowik, "Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification[Formula presented]," *Knowledge-Based Systems*, vol. 227, p. 107218, Sep. 2021, doi: 10.1016/j.knosys.2021.107218.
- [3] M. Dash and H. Liu, "Feature selection for classification," *Intelligent Data Analysis*, vol. 1, no. 3, pp. 131–156, 1997, doi: 10.3233/IDA-1997-1302.
- [4] M. Labani, P. Moradi, and M. Jalili, "A multi-objective genetic algorithm for text feature selection using the relative discriminative criterion," *Expert Systems with Applications*, vol. 149, p. 113276, Jul. 2020, doi: 10.1016/j.eswa.2020.113276.
- [5] B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimization for feature selection in classification: A multi-objective approach," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1656–1671, Dec. 2013, doi: 10.1109/TSMCB.2012.2227469.
- [6] R. Vijayanand and D. Devaraj, "A novel feature selection method using whale optimization algorithm and genetic operators for intrusion detection system in wireless mesh network," *IEEE Access*, vol. 8, pp. 56847–56854, 2020, doi: 10.1109/ACCESS.2020.2978035.
- [7] P. Bermejo, J. A. Gámez, and J. M. Puerta, "Speeding up incremental wrapper feature subset selection with Naive Bayes classifier," *Knowledge-Based Systems*, vol. 55, pp. 140–147, Jan. 2014, doi: 10.1016/j.knosys.2013.10.016.
- [8] A. Da Li, B. Xue, and M. Zhang, "Multi-objective feature selection using hybridization of a genetic algorithm and direct multisearch for key quality characteristic selection," *Information Sciences*, vol. 523, pp. 245–265, Jun. 2020, doi: 10.1016/j.ins.2020.03.032.
- [9] H. Banka and S. Dara, "A hamming distance based binary particle swarm optimization (HDBPSO) algorithm for high dimensional feature selection, classification and validation," *Pattern Recognition Letters*, vol. 52, pp. 94–100, Jan. 2015, doi: 10.1016/j.patrec.2014.10.007.
- [10] X. Zhang, Q. Zhang, M. Chen, Y. Sun, X. Qin, and H. Li, "A two-stage feature selection and intelligent fault diagnosis method for rotating machinery using hybrid filter and wrapper method," *Neurocomputing*, vol. 275, pp. 2426–2439, Jan. 2018, doi: 10.1016/j.neucom.2017.11.016.
- [11] M. Han and W. Ren, "Global mutual information-based feature selection approach using single-objective and multi-objective optimization," *Neurocomputing*, vol. 168, pp. 47–54, Nov. 2015, doi: 10.1016/j.neucom.2015.06.016.
- [12] B. Xue, W. Fu, and M. Zhang, "Multi-objective feature selection in classification: a differential evolution approach," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8886, 2014, pp. 516–528.
- [13] F. A. Hashim, E. H. Houssein, K. Hussain, M. S. Mabrouk, and W. Al-Atabany, "Honey badger algorithm: new metaheuristic algorithm for solving optimization problems," *Mathematics and Computers in Simulation*, vol. 192, pp. 84–110, Feb. 2022, doi: 10.1016/j.matcom.2021.08.013.





- [14] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: improving the strength pareto evolutionary algorithm," *Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems*, pp. 95–100, 2001, doi: 10.1.1.28.7571.
- [15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr. 2002, doi: 10.1109/4235.996017.
- [16] J. Knowles and D. Corne, "The Pareto archived evolution strategy: a new baseline algorithm for Pareto multiobjective optimisation," in *Proceedings of the 1999 Congress on Evolutionary Computation, CEC 1999*, 1999, vol. 1, pp. 98–105, doi: 10.1109/CEC.1999.781913.
- [17] Y. Liu and B. Cao, "A novel ant colony optimization algorithm with levy flight," *IEEE Access*, vol. 8, pp. 67205–67213, 2020, doi: 10.1109/ACCESS.2020.2985498.
- [18] D. Paul, A. Jain, S. Saha, and J. Mathew, "Multi-objective PSO based online feature selection for multi-label classification," *Knowledge-Based Systems*, vol. 222, p. 106966, Jun. 2021, doi: 10.1016/j.knsys.2021.106966.
- [19] A. M. Usman, U. K. Yusof, and S. Naim, "Filter-based multi-objective feature selection using NSGA III and cuckoo optimization algorithm," *IEEE Access*, vol. 8, pp. 76333–76356, 2020, doi: 10.1109/ACCESS.2020.2987057.
- [20] Y. Zhang, D. wei Gong, X. zhi Gao, T. Tian, and X. yan Sun, "Binary differential evolution with self-learning for multi-objective feature selection," *Information Sciences*, vol. 507, pp. 67–85, Jan. 2020, doi: 10.1016/j.ins.2019.08.040.
- [21] N. Kozodoi, S. Lessmann, K. Papakonstantinou, Y. Gatsoulis, and B. Baesens, "A multi-objective approach for profit-driven feature selection in credit scoring," *Decision Support Systems*, vol. 120, pp. 106–117, May 2019, doi: 10.1016/j.dss.2019.03.011.
- [22] J. González, J. Ortega, M. Damas, P. Martín-Smith, and J. Q. Gan, "A new multi-objective wrapper method for feature selection – Accuracy and stability analysis for BCI," *Neurocomputing*, vol. 333, pp. 407–418, Mar. 2019, doi: 10.1016/j.neucom.2019.01.017.
- [23] A. Sharma and R. Rani, "C-HMOSHSSA: Gene selection for cancer classification using multi-objective meta-heuristic and machine learning methods," *Computer Methods and Programs in Biomedicine*, vol. 178, pp. 219–235, Sep. 2019, doi: 10.1016/j.cmpb.2019.06.029.
- [24] H. E. Kiziloğlu, A. Deniz, T. Dokeroglu, and A. Cosar, "Novel multiobjective TLBO algorithms for the feature subset selection problem," *Neurocomputing*, vol. 306, pp. 94–107, Sep. 2018, doi: 10.1016/j.neucom.2018.04.020.
- [25] M. Amoozegar and B. Minaei-Bidgoli, "Optimizing multi-objective PSO based feature selection method using a feature elitism mechanism," *Expert Systems with Applications*, vol. 113, pp. 499–514, Dec. 2018, doi: 10.1016/j.eswa.2018.07.013.
- [26] E. Hancer, B. Xue, M. Zhang, D. Karaboga, and B. Akay, "Pareto front feature selection based on artificial bee colony optimization," *Information Sciences*, vol. 422, pp. 462–479, Jan. 2018, doi: 10.1016/j.ins.2017.09.028.
- [27] M. Dashtban, M. Balafar, and P. Suravajhala, "Gene selection for tumor classification using a novel bio-inspired multi-objective approach," *Genomics*, vol. 110, no. 1, pp. 10–17, Jan. 2018, doi: 10.1016/j.ygeno.2017.07.010.
- [28] C. M. Lai, "Multi-objective simplified swarm optimization with weighting scheme for gene selection," *Applied Soft Computing Journal*, vol. 65, pp. 58–68, Apr. 2018, doi: 10.1016/j.asoc.2017.12.049.
- [29] S. K. Nayak, P. K. Rout, A. K. Jagadev, and T. Swarnkar, "Elitism based Multi-Objective Differential Evolution for feature selection: A filter approach with an efficient redundancy measure," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 2, pp. 174–187, Feb. 2020, doi: 10.1016/j.jksuci.2017.08.001.
- [30] X. Z. Gao, M. S. R. Nalluri, K. Kannan, and D. Sinharoy, "Multi-objective optimization of feature selection using hybrid cat swarm optimization," *Science China Technological Sciences*, vol. 64, no. 3, pp. 508–520, Mar. 2021, doi: 10.1007/s11431-019-1607-7.
- [31] M. Rostami, S. Forouzandeh, K. Berahmand, and M. Soltani, "Integration of multi-objective PSO based feature selection and node centrality for medical datasets," *Genomics*, vol. 112, no. 6, pp. 4370–4384, Nov. 2020, doi: 10.1016/j.ygeno.2020.07.027.
- [32] B. Abdollahzadeh and F. S. Gharehchopogh, "A multi-objective optimization algorithm for feature selection problems," *Engineering with Computers*, vol. 38, no. S3, pp. 1845–1863, Aug. 2022, doi: 10.1007/s00366-021-01369-9.
- [33] A. Frank and A. Asuncion, "UCI machine learning repository," 2010, [Online]. Available: <http://archive.ics.uci.edu/ml>.

BIOGRAPHIES OF AUTHORS



Anusha Papasani     is full-time research Scholar in the School of Computing Science and Engineering at Vellore Institute of Technology -AP, India. She holds an M.Tech. Degree from Lakireddy Bali Reddy College of Engineering, JNTUK University. Her research interests lie in the area of evolutionary computation, soft computing, optimization techniques, and machine learning. She can be contacted at email: anoosha.papasani@gmail.com.



Nagaraju Devarakonda     is working as Associate Professor in the School of Computing Science and Engineering at VIT-AP, India. He has got totally 16 years of teaching experience in various engineering institutions in Andhra Pradesh. His research interests lie in the area of data mining, machine learning, soft computing, and optimization techniques. He can be contacted at email: dnagaraj_dnr@yahoo.co.in.