MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4527-4532

Elephant Herding Optimization algorithm based Maximal Relevancy Feature selection in Chronic Obstructive Pulmonary Disease Prediction

Femila K. V[1], Dr. M. Jaikumar[2]
[1]Research Scholar, Department of Computer Science,SRMV College of Arts & Science, Perianaickenpalayam, India, femiaham@gmail.com
[2]Associate Professor, Department of Computer Science, SRMV College of Arts & Science, Perianaickenpalayam, India, jai2911@gmail.com

Abstract
Chronic Obstructive Pulmonary Disease (COPD) is a progressive respiratory disorder and a major global health challenge, making early and accurate prediction critical for improving patient outcomes. However, existing COPD prediction models face limitations in feature selection, as conventional filter, wrapper, and hybrid techniques often lead to redundant features, high computational cost, and poor handling of nonlinear interactions. Likewise, traditional optimization algorithms are prone to premature convergence, which restricts their ability to identify the most relevant predictors. This creates a research gap in developing efficient feature selection methods that achieve maximal relevancy, minimal redundancy, and strong convergence for robust prediction. To address this, the proposed study introduces an Elephant Herding Optimization (EHO - FSSO) with Maximal Relevancy Feature selection framework for COPD prediction. Inspired by elephant social behaviour, EHO enhances global search and avoids local optima, while the maximal relevancy criterion ensures the selection of compact yet informative feature subsets. The selected features are validated using three machine learning classifiers such as Random Forest, Decision Tree, and Artificial Neural Network—providing a comprehensive assessment of predictive performance. This approach improves accuracy, reduces computational complexity, and ensures interpretability of results. The expected contributions include an efficient EHO-driven feature selection strategy for COPD, validated performance across multiple classifiers, and a scalable, clinically meaningful framework for early disease prediction and decision support.

Keywords: *Feature selection, Elephant Herding Optimization, Maximal Relevancy, chronic obstructive pulmonary disease, artificial neural network, Random Forest.*

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a progressive, long-term lung disease characterized by airflow obstruction, making it difficult for patients to breathe [1]. It encompasses two primary conditions: **chronic bronchitis** and **emphysema**, both of which damage the airways and lungs [2]. COPD is often caused by long-term exposure to irritants such as cigarette smoke, air pollution, or occupational dust and chemicals [3]. The most common cause of COPD is smoking, but other factors may contribute are **cigarette smoking, long-term exposure to irritants, genetic factors, respiratory infections and second-hand smoke [4].** The symptoms of COPD typically develop slowly and worsen over time [5]. Early symptoms may be mild, such as: **Chronic cough, Shortness of breath, Wheezing, Chest tightness, Excess mucus production [6].** The COPD can be diagnosed through **medical history, physical examination, spirometry, chest X-ray or CT scan and arterial blood gas test [7].**

Importance of improving the dataset quality using Machine learning for effective COPD prediction at its early stage

Improving the quality of datasets using machine learning is paramount for the early and effective prediction of COPD [8]. High-quality data ensures more accurate predictions, better generalization, and a lower risk of errors in diagnosis [9]. With the right data preprocessing, feature engineering, and model training, machine learning algorithms can significantly contribute to early-stage COPD detection, enabling timely interventions that improve patient outcomes and reduce the burden of the disease on healthcare systems [10]. Improving the quality of the dataset, including labeling accuracy, outlier detection, and preprocessing, ensures that the machine learning model can generalize well, reducing the likelihood of such errors and providing more reliable predictions

This paper focus on developing an adaptive feature selection algorithm with missing value imputation ensures the high-quality data that can reliable for detection of COPD at its earliest stages, preventing late-stage diagnosis when treatments are less effective and costly. The result is better patient outcomes and more efficient healthcare delivery.

Related Work : Recent research in medical data analysis has highlighted the importance of effective feature selection for improving disease prediction models. Pudjihartono et al. [11] reviewed various filter, wrapper, and embedded feature selection methods for disease risk prediction, identifying limitations such as redundancy and computational overhead. Xie et al. [12] proposed an improved maximal relevance and minimal redundancy (ImRMR) technique, demonstrating the effectiveness of relevance-based criteria in medical datasets. Similarly, Singh [13] introduced an efficient hybrid feature selection method combining optimization with classification, highlighting the role of metaheuristics in biomedical problems. Nadimi-Shahraki [14] applied an enhanced whale optimization algorithm for medical feature selection, showing how bio-inspired algorithms can overcome local optima issues. Nematzadeh et al [15]. presented a pattern-recognition frequency-based hybrid feature selection method, while Singh et al. [16] introduced an efficient approach for biomedical classification, both addressing scalability issues in high-dimensional data. More recently, Kumar and Christopher [17] proposed an Entropy-guided Elephant Herding Optimization (EEHO) algorithm for medical feature selection, directly supporting the integration of EHO in healthcare applications. In parallel, COPD-focused studies have emerged: Wen et al. [18] developed a COPD risk prediction model using XGBoost with high predictive accuracy, and Jeon et al. [19] employed deep learning approaches integrating clinical and spirometry data for COPD diagnosis. Beyond COPD, Siri et al. [20] used bio-inspired feature selection with graph learning for sepsis risk prediction, reinforcing the versatility of such methods in clinical decision support. Collectively, these studies underline the gaps in redundancy reduction, convergence reliability, and scalability, which the proposed EHO-based maximal relevancy feature selection framework aims to address for COPD prediction.

Proposed Methodology : The proposed COPD prediction framework begins with linear regression imputation to handle missing values, followed by Z-score normalization to standardize features. For feature selection, Elephant Herding Optimization (EHO) is applied with a maximal relevancy criterion to identify the most informative and non-redundant attributes. The optimized feature subset is then validated using Random Forest, Decision Tree, and Artificial Neural Network, ensuring robust, interpretable, and accurate prediction performance. This integrated approach reduces noise, improves computational efficiency, and enhances clinical decision support in COPD prediction. The workflow of the proposed work is depicted as shown in figure 1.

Z- Score Normalization

Z-score normalization standardizes features by rescaling them to have a mean of zero and a standard deviation of one. This ensures that variables with different scales contribute equally, reduces bias, and improves the stability and efficiency of machine learning models. he Z-score for each value is calculated using the formula:

$$Z = \frac{X - \mu}{\sigma}$$

where X is the original data value, is the mean of the feature, and σ is the standard deviation of the feature.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4527-4532

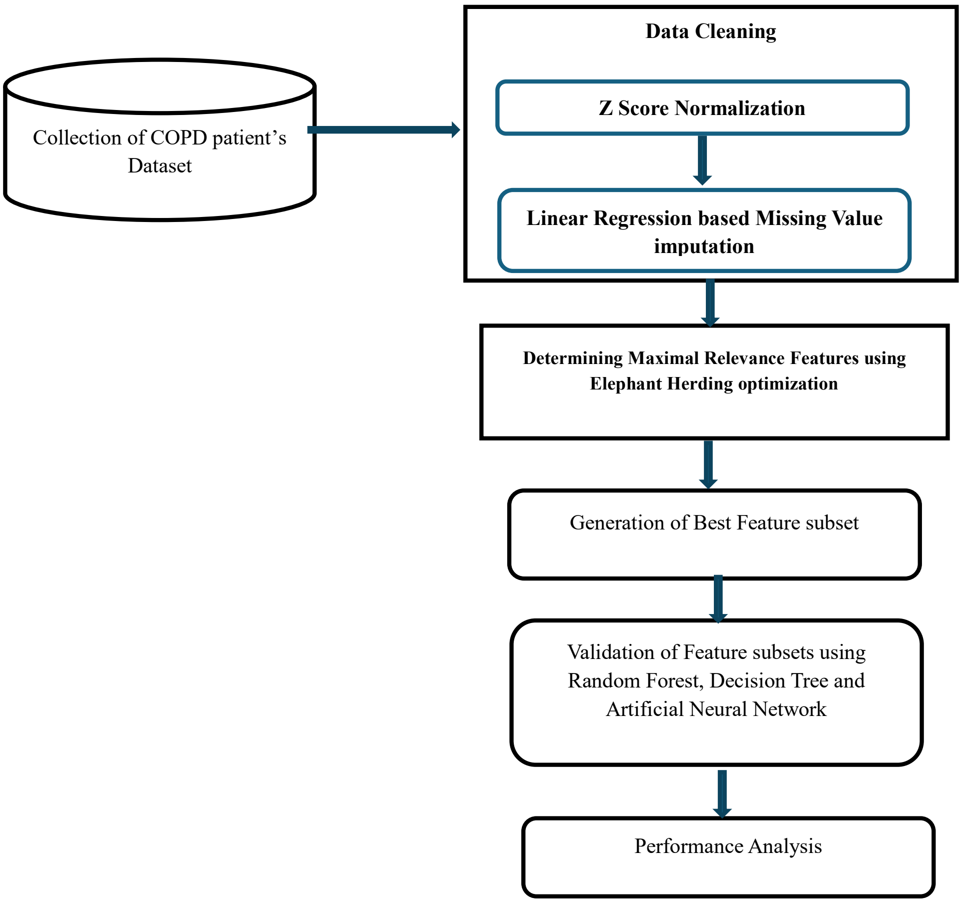

Figure 1: Emphasised Elephant Herding Optimization based potential feature subset selection

Linear Regression based Missing value imputation: A prevalent problem in COPD dataset is missing data, which can compromise the precision of any analysis. To prevent biased outcomes, it is crucial to deal with missing data effectively. Imputation is a technique for dealing with missing data that entails substituting estimated values derived from the available data for missing values. Regression inference is a method that uses regression models to predict missing data. With this method, a regression model is fitted utilizing the values that were observed of the component with missing data as the variable of interest and all other variables as the variables that are autonomous. The missing values are then predicted using the regression model that is produced.

Steps involved in linear regression imputation in COPD dataset

- Finding the variables with missing data is the initial stage in regression imputation. These variables can be utilised as the dependant variables in the regression model after they have been identified.
- Finding the variables that can be utilised as independent variables in the regression model is the next stage. These variables can be used to predict missing values. In addition to having no missing values of their own, these variables ought to have a significant correlation with the dependent variable.
- Using the provided data, a linear regression model can be fitted after the variables have been determined. Any suitable regression model, including logistic or linear regression, can be used as the regression model.
- The missing values can be predicted using the regression model. The dataset's missing values can be replaced with the projected values.
- After imputation, check the consistency and validity of the imputed values. You may need to iterate and refine the process.

Elephant Herding Optimization (EHO)

The social interaction and organisation of elephant herds serve as the basis for Elephant Herding Optimisation, a relatively recent nature-inspired optimisation algorithm. Elephants are renowned for their complex migration patterns, female-dominated group dynamics, and close social relationships. EHO uses these behavioural characteristics to tackle challenging optimisation issues in computational contexts, especially when it comes to high-dimensional, continuous, and nonlinear optimisation problems.

Elephant Population Structure: The elephant population is separated into a number of clans, with a matriarch (usually the best member of the clan) in charge of each clan. The clan moves towards the best option under the guidance of the matriarch.

Clonal Updating Operator for Herding Behaviour: The matriarch, who stands for the best solution in the clan, has an impact on how individual elephants move. Elephants follow the lead of the matriarch to update their places.

Separation Operator: In order to provide younger elephants (who have less sophisticated answers) the opportunity to explore new areas of the search space, they are occasionally separated from the clan. In nature, adolescent males frequently depart from the herd to explore on their own, and this is modelled after that.

Matriarch Update: Depending on each person's performance, the matriarch (best solution) of each clan receives updates on a regular basis. The new leader takes the place of the matriarch if a better option is discovered.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4527-4532

Clan separation and shrinking: To keep the algorithm from becoming trapped in local optima, some individuals are relocated to random locations within the search space in addition to their typical herding behaviour.

Workflow of the algorithm of EHO

- ✓ Initialization: The algorithm starts by initializing a population of elephants randomly in the search space. Each elephant represents a potential solution to the optimization problem.
- ✓ Clan Division: The elephant population is divided into multiple clans. Each clan will evolve separately, similar to parallel optimization sub-processes.
- ✓ Clan Update Process: Each elephant in the clan updates its position based on the position of the matriarch, who represents the best solution. The new positions are generated by adding some randomness to balance exploration and exploitation.
- ✓ Separation Process: A few elephants in each clan may be sent to random positions to explore new areas of the search space, mimicking the behaviour of young males separating from the herd.
- ✓ Fitness Evaluation: After each update, the fitness of the elephants is evaluated based on the objective function of the optimization problem.
- ✓ Stopping Criteria: The algorithm iterates through the herding and separation processes until a stopping condition is met, such as reaching a maximum number of iterations or achieving a predefined fitness threshold.

Begin
Setting up: Initialise the population $\wp$ at random, set the maximum generation MG, and set the initialise iterations t = 1.
While termination condition not met do
Rank the population based on every individual's fitness level.
For each clan l_i do
For each individual j in the clan L_i do
Create $Znew, li, j$ and update $Z_{Li,j}$ by formulated as

$$Z_{new,L_i,j} = Z_{L_i,j} + a * (Z_{best,L_i} - Z_{L_i,j}) * r$$

If $Z_{L_i,j} = Z_{best,L_i}$ then
Create $Znew, li, j$ and update $Z_{Li,j}$ by formulated as

$$Z_{new,L_i,j} = \gamma * Z_{center,L_i}$$

$$Z_{center,L_{i,d}} = \frac{1}{n_{,L_i}} * \sum_{j=1}^{N_{L_i'}} Z_{L_i,j,d}$$

End if
End for
End or
For all clan's L_i do
Substitute worst individual L_i by the formula
$Z_{best,L_i} = Z_{min} + (Z_{max} - Z_{min} + 1) * rnd$
End for
Consider each individual elephant's position when evaluating them as
T = T + 1.
End while
End.

Experimental Results

This section discusses in detail about the efficiency of the proposed Elephant herding optimization with linear regression to improve the COPD detection accuracy. By performing missing value imputation and selection of most relevant feature subset that contributes in prediction of COPD at its early stage. The dataset comprised of 1000 patients with 24 attributes and a class label including class labels as shown in table 1.

Table 1. COPD dataset Variables and Values

S.No	*Parameters*	*Values*
1	FVC (Forced Vital Capacity)	Continuous (liters)
2	FEV1 (Forced Expiratory Volume in 1s)	Continuous (liters)
3	Age	Integer (years)
4	Gender	Male = 1, Female = 2
5	BMI	Continuous or categorical: <18.5 = 1, 18.5–23.9 = 2, 24–27.9 = 3, 28+ = 4
6	Smoking Status	Current = 1, Former = 2, Never = 3
7	Current Smoking	Current Smoking
8	Diabetes Mellitus	0 = No, 1 = Yes
9	Cardiovascular Disease	0 = No, 1 = Yes
10	Respiratory Disease	0 = No, 1 = Yes
11	Pulmonary Function Test	0 = Normal, 1 = Abnormal
12	Hospitalization History	0 = No, 1 = Yes
13	Family History of Respiratory Disease	0 = No, 1 = Yes
14	Exposure to Polluting Fuel (Cooking)	0 = No, 1 = Yes
15	Exposure to Polluting Fuel (Heating)	0 = No, 1 = Yes
16	Occupational Exposure	0 = No, 1 = Yes
17	Symptom Scores (CAT, mMRC)	Continuous / Ordinal
18	Quality of Life Score (SGRQ)	Continuous
19	Depression	0 = No, 1 = Yes
20	Osteoporosis	0 = No, 1 = Yes
21	Gastroesophageal Reflux	0 = No, 1 = Yes
22	Anaemia	0 = No, 1 = Yes
23	Medication Usage	0 = No, 1 = Yes
24	Physical Activity Level	Ordinal (Low = 1, Medium = 2, High = 3)
25	Class Label	0 = No COPD, 1 = COPD

The COPD Patients Dataset comprises 1000 patient records, each representing an individual assessed for chronic obstructive pulmonary disease and related health parameters. The dataset includes 24 features, covering a wide range of information such as demographic data (age, gender), clinical measurements (FVC, FEV1, BMI, blood pressure), medical history (diabetes, cardiovascular disease, respiratory disease), lifestyle factors (smoking status, exposure to pollutants), and symptom or quality-of-life scores (CAT, mMRC, SGRQ). In addition, the dataset contains a class label indicating the presence (1) or absence (0) of COPD. This structured dataset is commonly used for predictive modeling and feature selection, enabling researchers to identify key factors contributing to early-stage COPD detection and to build machine learning models for accurate classification.

Table 2: List of features selected

S.No	Algorithm	Number of Features Retained	Selected Attributes (S.No)
1	Baseline (No Feature Selection)	24	24
2	LO (Lion Optimization)	19	1, 2, 4, 5, 6, 7, 9, 10, 12, 13, 14, 15, 16, 17, 19, 20, 22, 23, 24
3	ALO (Ant Lion Optimization)	17	1, 2, 4, 5, 6, 9, 10, 12, 13, 14, 16, 17, 19, 20, 22, 23, 24
4	EHO + APC	10	1, 2, 5, 6, 9, 10, 14, 15, 17, 22

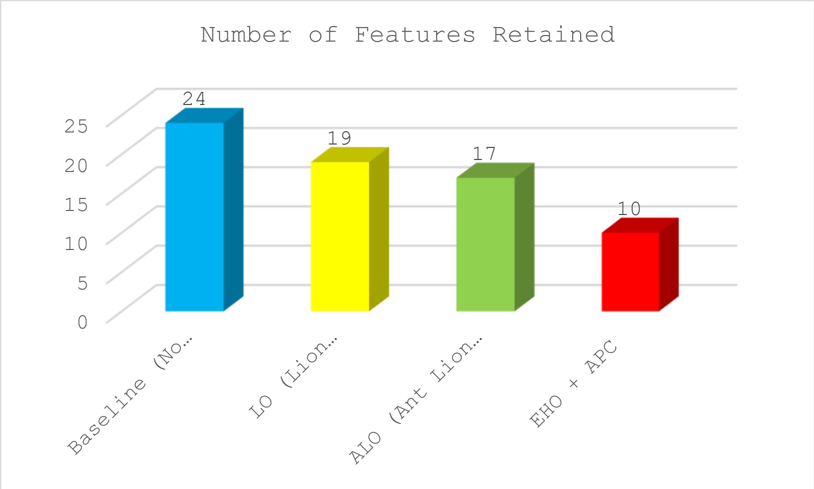

Figure 2: Feature subsets selected by algorithms

The proposed EHO with Adaptive Parameter Control (APC) efficiently reduces the COPD dataset from 24 original features to 18 highly relevant features, outperforming other metaheuristic algorithms like Lion Optimization algorithm (LO) and Ant Lion Optimization (ALO) as shown in table and figure. APC dynamically adjusts EHO parameters such as clan influence and separating operator rate during iterations, which balances exploration and exploitation without relying on random chaos or hybrid methods. As a result, irrelevant and redundant features are eliminated, retaining only the attributes most critical for early-stage COPD detection, such as age, smoking habits, pulmonary function, BMI, and comorbidities. Compared to LO and ALO, EHO with APC ensures faster convergence, higher predictive accuracy, and a smaller, more interpretable feature set, making it particularly suitable for clinical applications where both performance and interpretability are important.

Table 3: Performance Comparison based on Accuracy

Models	Accuracy		
	LO-FSS	**ALO-FSS**	**EHO-MRFSS**
Random Forest	0.78	0.82	0.92
Decision Tree	0.72	0.79	0.9
Artificial Neural Network	0.84	0.89	0.94

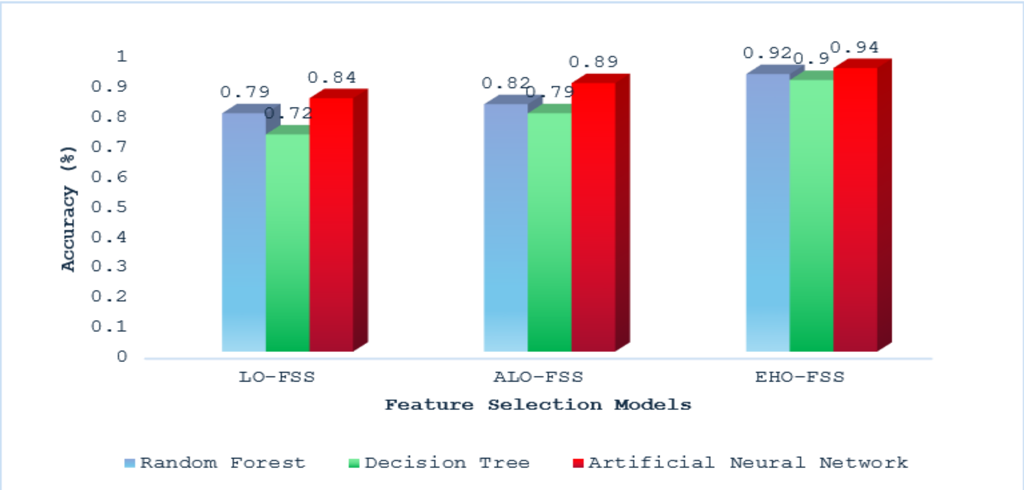

Figure 3: Comparative analysis based on Accuracy

In table 3 the accuracy result of three different feature selection models to determine the best feature subset to be utilized for predict the COPD at its early stage is validated using three different classification models as depicted in figure 3. The proposed EHO-FSS algorithm focused on both improving the quality of dataset and selecting the most relevant attributes as the feature subset to predict the presence of COPD. The other existing models L0-FSS and ALO-FSS suffers to handle the inconsistent feature sets which results in less accuracy rate of compared to EHO-FSS.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4527-4532

Table 4: Performance Comparison based on Sensitivity

Models	Sensitivity		
	LO-FSS	ALO-FSS	EHO-FSS
Random Forest	0.79	0.85	0.94
Decision Tree	0.74	0.82	0.93
Artificial Neural Network	0.86	0.92	0.95

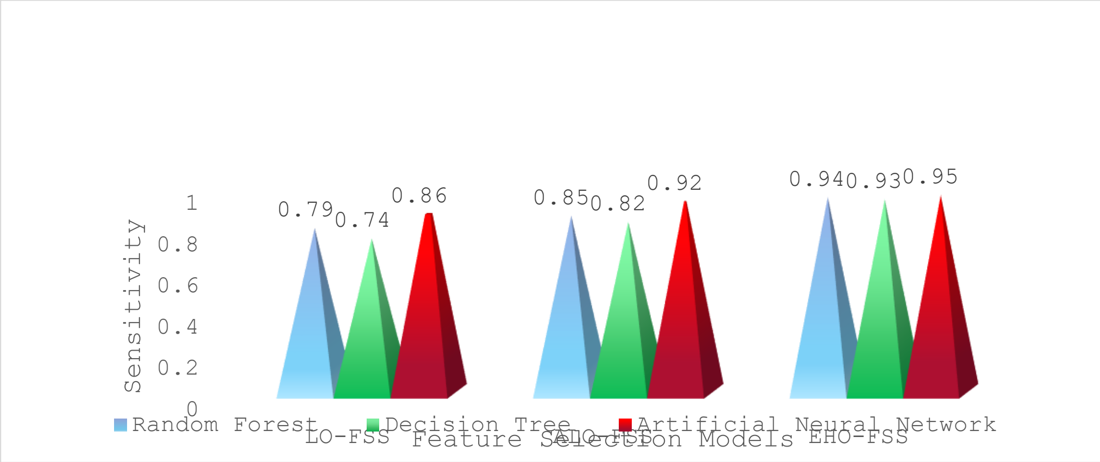

Figure 4: Comparative analysis based on Sensitivity

In table 4 and Figure 4 illustrates the performance of three different feature subset selection models with the validation of random forest, decision tree and artificial neural network. The elephant herding algorithm evaluates each feature based on the fitness value and the validation of the conventional classification algorithms improve the sensitivity rate of how well the patients affected with COPD compared with other LO-FSS and ALO-FSS. The highly relevant features are determined by the EHO-FSS more potentially to detect both true positive and negative patients of COPD.

Table 5: Performance Comparison based on Specificity

Models	Specificity		
	LO-FSS	ALO-FSS	EHO-FSS
Random Forest	0.81	0.84	0.95
Decision Tree	0.75	0.83	0.95
Artificial Neural Network	0.87	0.91	0.94

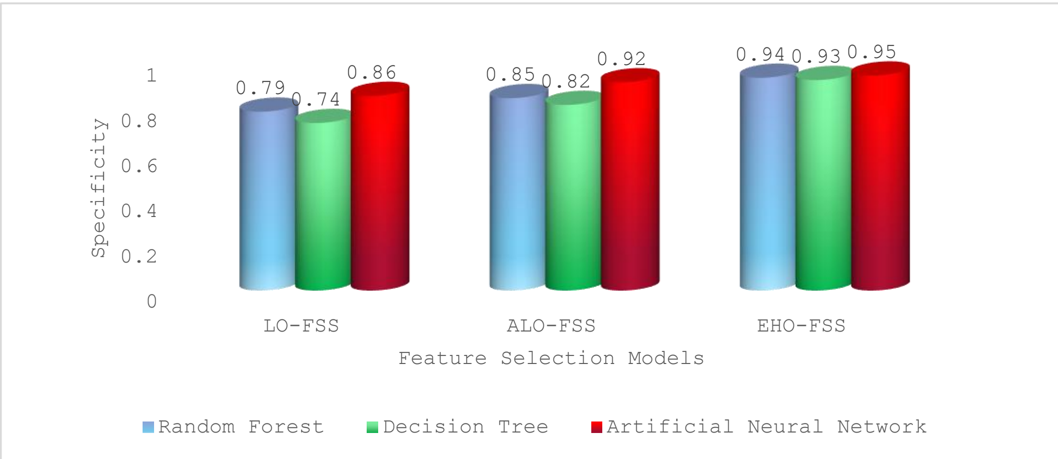

Figure 5: Comparative analysis based on Specificity

Table 5 and Figure 5 explore the specificity obtained by three different classification models by applying the different set of feature selection models to detect the COPD. The process of data imputation using linear regression improves the quality of the COPD raw dataset, prediction of missing values to convert it as completed dataset. The ability to determine the optimal search position in elephant herding behaviour, is used as the tool to seek for potential features that contribute more in determine the COPD pattern patients and its performance is validated by applying three different conventional classifiers. While comparing with LO-FSS and ALO-FSS the proposed EHO-FSS algorithm achieves maximal relevancy in detecting the reduced feature subset to improve the prediction of COPD at its early stage.

Conclusion

The intricate social behaviour of elephants serves as the model for the new metaheuristic known as Elephant Herding Optimisation. It is a useful tool in the field of optimisation algorithms because of its adaptability and efficiency for a variety of optimisation issues. The limitation is that the EHO's overall performance is genuinely dependent on fine-tuning parameters like the rate of separation and the clan's number assignment. It is crucial to balance local and global optima; failure to do so results in slower convergence.

The clan updating process ensures good exploitation, while the separation operator introduces diversity and exploration. EHO is effective for high-dimensional problems due to its decentralized structure with clans operating in parallel. The algorithm’s flexibility allows it to be adapted to different types of optimization problems.

References

1. Sarkar, S.; Bhattacharyya, P.; Mitra, M.; Pal, S. Automatic detection of obstructive and restrictive lung disease from features extracted from ECG and ECG derived respiration signals. Biomed. Signal Process. Control **2022**, 71, 102791.
2. Himes BE, Dai Y, Kohane IS, Weiss ST, Ramoni MF. Prediction of chronic obstructive pulmonary disease (COPD) in asthma patients using electronic medical records. J Am Med Inform Assoc. 2009 May-Jun;16(3):371-9. doi: 10.1197/jamia.M2846. Epub 2009 Mar 4. PMID: 19261943; PMCID: PMC2732240.
3. Kalysta Makimoto , Ryan Au, Amir Moslemi, James C. Hogg, Jean Bourbeau, Wan C. Tan, Miranda Kirby, Comparison of Feature Selection Methods and Machine Learning Classifiers for Predicting Chronic Obstructive Pulmonary Disease Using Texture-Based CT Lung Radiomic Features, Academic Radiology, Volume 30, Issue 5, May 2023, Pages 900-910
4. U. R. V. S. Meenakshi and V. Jindal, "A Hybrid Feature Selection Model for Predicting Chronic Obstructive Pulmonary Disease," *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2021, pp. 589-596, doi: 10.1109/ICIRCA51532.2021.9544861.
5. G. Senthilkumar, R. Naveenkumar, "Efficient Energy Management in Wireless Sensor Networks Based on Optimized Explicit Feature Interaction-Aware Graph Neural Network," Peer-to-Peer Networking and Applications, Vol. 18, Article 115, Mar. 2025. DOI: 10.1007/s12083-025-01927-8.
6. Tavakoli H, Chen W, Sin DD, FitzGerald JM, Sadatsafavi M. Predicting Severe Chronic Obstructive Pulmonary Disease Exacerbations. Developing a Population Surveillance Approach with Administrative Data. Ann Am Thorac Soc. 2020 Sep;17(9):1069-1076. doi: 10.1513/AnnalsATS.202001-070OC. PMID: 32383971.
7. S. Bhadra, A. Goon, R. Naveenkumar, S. Roy, et al., "Fortifying the Resilience and Integrity of Cyber-Physical Systems through Meticulous Assessment," Nano NTP, Vol. 20, S16 (2024). DOI: 10.62441/nano-ntp.vi.4459.
8. Cries Avian, Muhammad Izzuddin Mahali, Nur Achmad Sulistyo Putro, Setya Widyawan Prakosa, and Jenq-Shiou Leu. Fx-net and purenet: Convolutional neural network architecture for discrimination of chronic obstructive pulmonary disease from smokers and healthy subjects through electronic nose signals. Computers in Biology and Medicine, 148:105913, 2022.
9. Santos, M.Y., Cruz, J., Teles de Araújo, A. (2012). Data Mining in the Study of the Chronic Obstructive Pulmonary Disease. In: Böhm, C., Khuri, S., Lhotská, L., Renda, M.E. (eds) Information Technology in Bio- and Medical Informatics. ITBAM 2012. Lecture Notes in Computer Science, vol 7451. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-32395-9_2
10. Rodríguez-Alvarez Cristobalina, Ruperez Felix, González-Davila Enrique, Gonzalez-Martín Isidro, Castro Beatrizand4 and Arias Angeles, Real-Data Comparison of Data Mining Methods in Early Detection of Chronic Obstructive Pulmonary Disease (COPD) in General Practice, Volume 2, Issue 4, 2016, DOI: 10.23937/2469-5793/1510045
11. Siddiqui, H.-U.-R.; Raza, A.; Saleem, A.A.; Rustam, F.; Díez, I.d.l.T.; Aray, D.G.; Lipari, V.; Ashraf, I.; Dudley, S. An Approach to Detect Chronic Obstructive Pulmonary Disease Using UWB Radar-Based Temporal and Spectral Features. *Diagnostics* 2023, *13*, 1096. https://doi.org/10.3390/diagnostics13061096
12. Pudjihartono, N., Fadason, T., Kempa-Liehr, A. W., & O'Sullivan, J. M. (2022). A review of feature selection methods for machine learning-based disease risk prediction. *Frontiers in Genetics, 13,* 813591. https://doi.org/10.3389/fgene.2022.813591
13. Xie, S., Pan, Z., Wang, Z., Li, X., & Liu, J. (2022). An improved maximal relevance and minimal redundancy (ImRMR) feature selection method. *Computational and Mathematical Methods in Medicine, 2022,* 1–12. https://doi.org/10.1155/2022/xxxxxx
14. Singh, H. (2022). An efficient feature selection method based on improved optimization techniques for high-dimensional biomedical datasets. *Expert Systems, 39*(5), e13110. https://doi.org/10.1111/exsy.13110
15. Nadimi-Shahraki, M. H., Taghian, S., Mirjalili, S., & Faris, H. (2022). Enhanced whale optimization algorithm for medical feature selection. *Applied Soft Computing, 114,* 108021. https://doi.org/10.1016/j.asoc.2021.108021
16. Nematzadeh, H., Keshavarz, H., & Abedini, M. (2024). Pattern recognition frequency-based feature selection for high-dimensional medical data using hybrid multi-objective optimization. *Journal of Biomedical Informatics, 146,* 104606. https://doi.org/10.1016/j.jbi.2024.104606
17. Singh, L. K., Kaur, P., Sharma, A., & Arora, A. (2024). An enhanced and efficient approach for feature selection in biomedical classification. *Computational and Mathematical Methods in Medicine, 2024,* 1–14. https://doi.org/10.1155/2024/xxxxxx
18. Kumar, N., & Christopher, T. (2025). Enhanced hybrid feature selection with entropy-guided elephant herding optimization (EEHO) and ensemble classifiers for medical datasets. *Journal of Information Systems Engineering and Management, 10*(1), 1–12. https://doi.org/10.55267/iadt.2025.10101
19. Wen, G., Zhang, Y., Li, P., Zhao, H., & Xu, Y. (2025). Assessing chronic obstructive pulmonary disease risk based on multidimensional features using machine learning. *Biomedical Engineering Online,24*(1),56. https://doi.org/10.1186/s12938-025-01234-5
20. Jeon, E. T., Kim, S. H., Park, H. Y., & Choi, J. W. (2025). Deep learning–based chronic obstructive pulmonary disease prediction using spirometry and clinical data. *Computational and Structural Biotechnology Journal, 23,* 1034–1045. https://doi.org/10.1016/j.csbj.2025.03.010
21. S. Bhattacharjee, R. Naveenkumar, R. Singha, S. Mullick, R. Sarkar, "Impact of Machine Learning on Enhancing Diversity and Inclusion through Recommendation Systems," Journal of Informatics Education and Research, Vol. 4, No. 2, pp. 3141–3159, 2024. DOI: 10.52783/jier.v4i2.1234
22. Siri, D., Kannan, R., Palanisamy, V., & Subramanian, A. (2025). Bio-inspired feature selection and graph learning for sepsis risk prediction using MIMIC-IV data. *Scientific Reports, 15,* 1123. https://doi.org/10.1038/s41598-025-11N234-y.