ISSN 2007-9737

Meta-Heuristic Hybrid Wrapper Method based on Feature Selection for Classification of Biological Samples

Sabita Rani Behera[1], Bibudhendu Pati[1,*], Sasmita Parida[2]

[1] Rama Devi Women's University,
Department of Computer Science,
India

[2] SoA Univesity,
Institute of Technical Education and Research,
India

sabita_behera77@yahoo.com, {patibibudhendu, sasmitaparida2004}@gmail.com

Abstract. Cancer is the vital cause of death across the Globe. Microarray technology is regarded as a promising diagnostic and classification tool for cancer. It explores genetic mutations occurring within a cancer cell. Dimensionality reduction techniques (DRT) are vital in microarray based data analysis. The microarray data contains a huge number of attributes or dimensions, which can adversely affect performance parameters of the model. Hence it is necessary to identify most relevant attributes to be retained and discard rest attributes. Statistical and machine learning (ML) techniques are employed to identify the majority of important genes or attributes to be retained. Two wrapper hybrid wrapper models are proposed for the feature selection purpose. The first hybrid method combines the Grey Wolf Optimization (GWO) with the Jaya optimization method, whereas second hybrid method combines GWO and Particle swarm optimization (PSO) algorithm. These two hybrid models are applied individually on four benchmark microarray datasets containing data on cancer of the central nervous system, Breast cancer, ovarian cancer and leukaemia cancer to get reduced datasets. Classification algorithms Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Naive Bayes (NV), and Linear Discriminant Analysis (LDA) are classification models used individually to classify malignant and benign genes from each category of reduced data sets with stratified 10-fold cross-validation. Classification accuracy of all classifiers on individual dataset for both wrapper hybrid models is compared with each other.

Keywords: Feature selection (FS), PSO, GA, Jaya, GWO.

1 Introduction

Cancer is a cell-based disease that unconditionally develops and spreads to other areas of the human body. Genes that regulate cell growth and division are altered in cancer. Cancerous tumors such as malignant tumors and non-cancerous tumors (benign tumors), can harm cells. Early detection and diagnosis are essential for the control of the disease and to increase the survival probability in the case of a malignant condition, which gradually progresses towards death [1,45]. The expression of the genes within a tumor must determine how it behaves. Therefore, it is vital to identify gene sets whose expression or absence identifies each unique characteristic of a tumor. A previously unknown and clinically significant tumor subgroup may be found using microarray technology [2].

Microarrays were used to investigate the potential relationship between gene expression changes and cancer treatment outcomes. Researchers can utilize microarray technology (MT) to analyze 10,000 genes' activities in one experiment and learn essential information about how cells function. This technique can produce gene expression data, which is critical for classifying and forecasting cancer disease. However, most of the high-dimensional data in the genes are unnecessary, redundant, and noisy, which are not helpful in the identification of diseases. The use huge number of features in

small sample sizes in the medial domain leads to poor diagnosis and overfitting in classification in ML [3, 4].

Dimensionality reduction is a highly stimulating field of study machine learning, statistics, and pattern recognition. Dimensionality reduction aims to improve the efficiency of the classification algorithm by omitting features from the Microarray dataset that are redundant or unnecessary. There are several methods for reducing dimensions. However, the domain of use and the uniqueness of the dataset dictate which dimensionality reduction approach should be used. FS and Feature extraction (FE) are subcategories of DRT [43, 46]. In FS the subset of original feature sets are selected that are appropriate for classification is known as FS. The process of choosing the best feature among those available involves two steps: (1) Finding subsets of the feature, and (2) Evaluating the subsets [5]. A new feature set is generated during feature extraction from the already existing feature sets. The FS methods are of the types such as filter, wrapper and embedding approaches [5, 6].

Filter techniques don't need a learning algorithm; instead, they operate on the inherent qualities of the data [43, 46]. Filter techniques are comparatively faster, but since FS is performed without consulting a learning system, filter methods typically yield worse accuracy results than wrapper methods. Relief, the Similarity measure, the Gini index, and so on are a few instances [7]. Wrapper approaches necessitate the use of a learning algorithm, which increases accuracy but also significantly increases computation time. PSO, genetic algorithms (GA), ACO etc. are some wrapper based FS methods used [8-10].

Filter and wrapper techniques are mostly applied together in the construction of embedded methods as a compromise between these. These methods attempt to combine intrinsic data qualities and learning algorithms into a method while striking a balance between the two kinds of methods. Precision and calculation time may be finely balanced, or it may be possible to reduce computing costs without sacrificing precision. The design of embedded systems has thus become the general tendency.

Using the conventional methods of FS, it is impossible to get the best subset from the available data sets. Therefore, to successfully minimize the dimensionality of microarrays, researchers have employed efficient hybrid algorithms to perform FS [43, 46].

1.1 Motivation and Contributions

Gene expressions in microarray data are interdependent. Subtle interaction among gene expressions often influences the outcome of the model. Traditional feature selection methods may fail to capture these intricate relationships. Traditional optimization techniques are not effective at identifying an optimal or near-optimal subset of features from large feature set. Researchers have developed new algorithms inspired by nature to address challenging optimization issues. ABC, BCOFS [19], PSO [9, 44], ACO [10], and CSA [11] are some popular optimization algorithms used. Investigating meta-heuristic algorithms motivates the researchers to solve challenging optimization issues for which Meta-heuristic algorithms are more effective for exploring large search space. These techniques are now recognized as some of the most useful ones for addressing a variety of real-world issues, such as gene selection [17,18]. A hybrid approach can give us global optimal solutions without stocking in local minima.

This work gives a thorough analysis of the meta-heuristic techniques for FS:

- To implement PSO, GWO for selecting the optimal set of features.
- To implement Jaya, GWO for selecting another optimal feature set.
- To implement simple classifiers SVM, DT, RF, NB and LDA individually on the optimized datasets for classification.
- Performance comparison is done between two models.

The subsequent text elucidates the organization of the article.

Some of the related works are summarized in Section 2. Section 3 presents a comprehensive explanation of the theory of gene selection methods. The categorization of gene selection using meta-heuristic strategies employed for

ISSN 2007-9737

Table 1. Summary of related works

Ref	Technique Applied	Problem Addressed	Model Evaluation Metric
[30]	Microarray data analysis, feature (gene) selection and genetic algorithm (GA)	Cancer classification for high-dimensional microarray data using bio-inspired methods	Accuracy
[31]	combination of Modified GWO with filter approach	To enhance gene selection for cancer classification	Accuracy
[32]	Hybrid feature selection approach combining 5 filters and one wrapper method	To enhance micro-array data analysis Top of Form	Classification accuracy
[33]	A hybrid wrapper approach named TLBOSA integrated with SA and TLBO techniques.	Improve the gene selection and handle local optima and exploitation limitations.	Accuracy
[34]	hybrid wrapper model named BTLBOGSA, TLBO and GSA	Enhancing gene selection accuracy for cancer class prediction	Accuracy
[35]	Independent Component Analysis (ICA), CS algorithm, ABC and GA	Classifying cancer types for imbalanced data distribution and low sample sizes in microarray datasets.	Accuracy

datasets with high dimensions is outlined in Section 4. The experimental findings from the study are presented in Section 5 using meta-heuristics-based FS techniques. Finally, section 6 concludes the findings with future scopes.

2 Related Works

Microarray technology is a prominent tool in genetic research, particularly in studying gene expression levels in organisms. The vast amount of data poses analytical challenges, especially in cancer classification, where FS is crucial due to high dimensionality, small sample sizes, and noise in data. Hambali *et. al*. presented a survey of FS methods in microarray cancer classification, addressing key challenges for open research issues [15].

Ghosh *et al.* [16] introduced a novel approach combining Ant Colony Optimization (ACO) with wrapper-filter methods for FS. He used a filter method for subset evaluation by which computational complexity can be reduced. Additionally, a memory mechanism and feature dimension-dependent pheromone update enhance the algorithm's multi-objective FS capability. Evaluation of diverse real-life datasets (UCI & NIPS2003 FS), using KNN and MLP classifiers, demonstrates superior performance compared to popular FS methods, particularly in emotion recognition and microarray datasets. Alirezanejad *et al.* focused on two heuristic gene selection methods: Xvariance and Mutual congestion measures. Evaluation of eight binary medical datasets demonstrates that Xvariance performs effectively with standard datasets, whereas Mutual Congestion significantly improves accuracy in high-dimensional datasets [17].

Goudos *et al.* proposed two antennas tailored for 3.7 and 26 GHz frequencies, developed using a novel evolutionary algorithm, GWO-Jaya. The algorithm, blending features from GWO and Jaya, is evaluated on benchmark functions and subsequently applied to antenna design. Performance assessment through fabrication and measurements demonstrates both antennas' circular polarization, wide-band behavior, and agreement between simulation and measurement results, affirming their efficacy [18].

Abdo *et al.* [19] designed a hybrid FS approach combining the chi-square filter method with GWO

and PSO. Evaluation of two datasets demonstrates significant accuracy improvements, with PSOGWO achieving a 95.3% boost and chi2-PSOGWO yielding 95.961% accuracy with time.

Dankolo *et al.* [20] developed an approach for FS and breast cancer classification, employing the FPA in conjunction with SVM on microarray data. Results indicate the effectiveness of FPA-SVM, surpassing the performance of Particle Swarm Optimization with an accuracy of 80.11%, demonstrating its promise for improved classification outcomes in breast cancer research.

Bhattacharyya *et al.* [21] proposed an efficient FS algorithm, Mayfly-Harmony Search (MHS), combining the Mayfly Algorithm (MA) and Harmony Search meta-heuristics. By integrating an S-shaped transfer function, the MA is adapted for binary FS.

Meenachi *et al*. introduced two novel FS algorithms, ACTFRO and GATFRO, which hybridize Tabu search with global optimal FS algorithms (ACO & GA). Aimed at cancer prediction from microarray gene expression data, these algorithms effectively balance global and local FS [22].

To select and classify genes across twelve high-dimensional cancer datasets, Hameed *et al.* [23] carried out a thorough comparative study of three nature-inspired meta-heuristic algorithms: binary-PSO, GA, and CS. The study uses a three-phase hybrid technique wherein five classification algorithms are evaluated in each step using Pearson product-moment correlation coefficient filtration. Results indicate that while BPSO demonstrates superior accuracy, CS exhibits efficiency by selecting fewer genes with reduced computational complexity compared to GA and BPSO. Table 1 summarizes some of the related works.

3 Feature Selection

In the procedure of identifying the most expressive genes meant for classification from a set of sample genes, FS methods and microarray data analysis are commonly employed. It is a process of choosing relevant features from the original dataset by eliminating inappropriate, irrelevant, or unnecessary features [30]. Filter, wrapper and embedding methods are the FS types, rely on how they work together to build the classification model [36, 43].

3.1 Filter Approach

Using its general statistical qualities, each feature is assessed separately in the filter technique [30]. The filter method does not make use of any particular learning model. It is therefore unaffected by the classifier. This method involved ranking the common features (genes) according to predetermined standards and then choosing the traits with the highest scores. Some popular filter methods are Information Gain (IG), Minimum Redundancy Maximum Relevance (mRmR)], and Chi2, Relief.

3.2 Wrapper Approach

Wrapper approaches utilize a classifier and learning algorithms to discover the most optimal subset of feature set. Wrapper methods entail exploring the space of primary features and getting a subset that aligns with the learning objectives. Notably, wrapper methods are associated with high computational costs and are often deemed unsuitable for high-dimensional datasets [1].

Nevertheless, their effectiveness surpasses that of feature-ranking algorithms, as they incorporate the classifier hypothesis. The following are descriptions of commonly employed wrapper methods categorized according to their meta-heuristic principles, as outlined in [1, 36].

The FS method often begins by partitioning the data into training data and validation data. The training data and test data ratio is taken as 90% and 10%. The chosen characteristics are utilized as input for the classifier, and the most favorable collection of features is determined based on the best-attained accuracy in classification.

The wrapper technique uses the classifier for training and sequentially evaluates the generalization performance of each relevant feature subset. This evaluation entails assessing the precision of the machine learning model that was trained using a specified subset of features when it is applied to the original dataset.

3.2.1 Meta-heuristic Approach

Meta-heuristic algorithms have been crucial in optimization due to the high cost of exhaustive search.

a) Genetic Algorithm

GA has many applications in various scientific and technical domains, including immune systems, economics, machine learning, optimization, ecology, population genetics, evolution and learning, and social systems [36]. Drawing influence from natural selection and evolution processes, GA functions as a heuristic search method. Three fundamental functions of the algorithm are crossover, mutation, and selection.

To eliminate those who are unfit for solving the current problem, the technique commences with the selection operation, wherein the most genetically superior people are chosen. After that, the next-generation process is applied to the chosen individuals. Subsequently, the crossover technique is executed, which creates new people by combining the selected ones. To favor the selection of the fittest, two individuals are randomly selected and their genes are exchanged. The mutation operation—which adds tiny, random modifications to the new solution (individuals)—concludes the algorithm [38].

b) Particle Swarm Optimization

PSO is based on the collective behavior observed in bird flocks. Similar to birds migrating in flocks toward a similar destination, the collaboration of the flock produces intelligence and efficiency [32, 34]. PSO exploits the motion of particles in an n-dimensional space to solve an optimization problem with n variables. The particles possess velocities that dictate their trajectory, while their fitness values are evaluated by the fitness function for optimization purposes. The particles move through the problem space as the best existing solutions follow them. PSO iterates through the issue space, updating each generation as it goes, beginning with a random set of particles, or solutions [46]. PSO is regarded as one of the better FS methods since it can efficiently scan large areas with little computational overhead. It also requires fewer parameters and is simpler to construct [32, 44].

This strategy successfully balances exploration and exploitation in the search space. The velocity of each particle is updated by using the equation 1:

$$V_{i,D}(t+1) = \omega * V_{i,D}(t) + \tau_1 * r_1 * \left(Pbest_{i,d} - x_{i,d}\right) + \tau_2 * r_2 * \left(Gbest_{i,d} - x_{i,d}\right), \quad (1)$$

$V_{i,D}(t+1)$ is the velocity of the particle I at the current time t+1 with dimension D, ω is the inertia weight, $\tau_1 \; and \; \tau_2$ is cognitive and social components, $r_1, and \; r_2$ are the random numbers from a distribution in the range of [0, 1]. $Pbest_{i,d} \; and \; Gbest_{i,d}$ are the local and global best of the particle i.

c) Grey Wolf Optimizer (GWO)

GWO is a meta-heuristic optimization algorithm, based on the social hierarchy and hunting behavior of grey wolves. The system works by using a group of wolves, where each wolf represents a possible solution to the optimization issue [33]. The system employs a hierarchical structure consisting of alpha, beta, delta, and omega wolves.

The GWO algorithm utilizes update equations to dynamically alter the placements of wolves, effectively balancing the exploration and exploitation aspects of the search process to efficiently explore the solution space. The GWO has 5 different phases including the Social Hierarchy, Encircling the prey, Hunting, Exploitation, and Exploration. The detailed work of GWO can be illustrated in Figure 1.

I. Social Hierarchy

Social hierarchy is a vital factor in the behavior and convergence of the GWO towards optimum solutions. The GWO algorithm draws inspiration from the social structure and hunting behavior of grey wolves, characterized by a distinct hierarchy within their packs. The social structure in GWO is delineated by four essential functions or places that wolves may assume within the pack:

Alpha Wolf (α):

– The alpha wolf symbolizes the most authoritative and top-ranking wolf in the group.

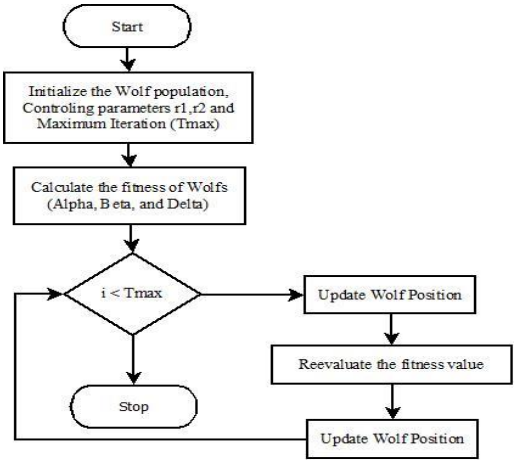

Fig. 1. GWO algorithm

- Within the context of GWO, the alpha wolf represents the most optimal solution that has been discovered so far throughout the process of optimization.
- The alpha wolf's position is adjusted according to the locations of the other wolves in the pack.

Beta Wolf (β):

- The beta wolf has the second highest position in the pack's hierarchy, placing just behind the alpha wolf in terms of dominance.
- Within the context of GWO, the beta wolf represents the second most optimal solution that is discovered throughout the process of optimization.
- The beta wolf's location is adjusted iteratively, similar to the alpha wolf, by taking into account the positions of other wolves.

Delta Wolf (δ):

- The delta wolf has the third highest position in the pack's social structure.
- Within the context of GWO, the term "delta wolf" refers to the third most optimal solution that is discovered throughout the process of optimization.
- Similar to the alpha and beta wolves, the location of the delta wolf is continuously adjusted as the algorithm advances.

Omega Wolf (ω):

- The omega wolf has the lowest position within the pack hierarchy and often symbolizes the least powerful or least desirable solution discovered.
- Within the context of GWO, the location of the omega wolf may not be updated as often as the positions of the other wolves due to its association with the least favorable option.
- Nevertheless, the hierarchical status of the omega wolf might still have an impact on the locomotion of other wolves within the pack.

The social hierarchy in GWO is used to direct the investigation and utilization of the search space. Wolves occupying higher-ranking locations (alpha, beta, and delta) have a more potent impact on the search direction. This is because their positions are used to update the placements of other wolves in later iterations. The hierarchical structure enables the algorithm to efficiently approach favorable sections of the solution space and simultaneously investigate unexplored areas for possible enhancements.

II. Encircling the Prey

Wolves navigate the search space by adjusting their locations according to the positions of other wolves and random variables.

The update of a wolf's location may be mathematically expressed by the following equation: $Xi + 1$ is the next position, Xi is the current position of wolf X, A is the controlling coefficient, and D is the direction vector between wolves and random variable.

Based on the prey Xp the next position of the wolve X can be determined as follows:

$$Xi + 1 = Xp - A \cdot Di, \tag{2}$$

$$Di = C.X - Xi, \tag{3}$$

$$A = 2 * a * r1, \tag{4}$$

$$C = 2 * r2, \tag{5}$$

where A is linearly decreased from 2 → 0, r1, and r2 are the random variable in the range of [0,1].

III. Hunting

In the Grey Wolf Optimizer (GWO), the hunting phase corresponds to the period during which wolves modify their locations by imitating the collaborative hunting behavior seen in nature among grey wolves, taking into account the positions of other wolves in the pack. The hunting process starts with wolves engaging in cooperative efforts to locate and pursue prey, aiming to find the most advantageous solutions. Each wolf adjusts its posture by imitating the social behaviors of a wolf pack, taking into account the locations of other wolves. The Alpha, Beta, Delta and Omega wolfs update their current position:

$$X_\alpha = X_\alpha - A \cdot D_\alpha , \quad (6)$$

$$X_\beta = X_\beta - A \cdot D_\beta , \quad (7)$$

$$X_\delta = X_\delta - A \cdot D_\delta , \quad (8)$$

$$D_\alpha = C_\alpha * X_\alpha - X, \quad (9)$$

$$D_\beta = C_\beta * X_\beta - X , \quad (10)$$

$$D_\delta = C_\delta * X_\delta - X . \quad (11)$$

IV. Exploitation Phase

The exploitation phase is a vital process when the algorithm systematically enhances and refines the best solutions found during the exploration phase. This phase aims to use the information acquired from earlier iterations to move towards more optimum solutions. During the exploitation phase, special focus is placed on wolves that hold higher ranks, such as the Alpha, Beta, and Delta wolves. These wolves are considered the best-performing solutions within the population. These wolves are crucial in directing the search towards potential areas of the solution space and enabling iterative improvement.

V. Exploration Phase

The exploration phase of the Grey Wolf Optimizer (GWO) is vital one as the algorithm actively investigates and explores the solution space to uncover probable optimum solutions. During this stage, the algorithm uses a varied array of tactics to encompass a broad spectrum of solutions and pinpoint favorable areas inside the search space. The wolf population, which includes Alpha, Beta, Delta, and Omega wolves, collectively participate in exploratory activities to find high-quality solutions. During the exploration phase, wolves adaptively modify their locations using precise update equations that take into account the positions of other wolves, random variables, and control settings. Wolves may navigate the solution space by dynamically moving, which helps them strike a balance between investigating new areas and making use of existing solutions. The exploration phase is essential for commencing the search process, broadening the exploration tactics, and establishing the groundwork for the following phases such as exploitation and convergence. The algorithm may uncover and recognize possible prey (ideal solutions) by conducting efficient exploration. This lays the foundation for subsequent phases of the optimization process when additional refining and improvement can take place.

Jaya Algorithm

The Jaya Optimization Algorithm is a method of optimizing a population by making improvements based on certain criteria. It draws inspiration from the notion of collective behavior, in which the whole population collectively works towards enhancing the greatest answer that has been discovered so far. Jaya, unlike other optimization methods, does not depend on probabilistic models or derivatives, making it a derivative-free optimization approach [26, 40]. The following is the elucidation of the Jaya Optimization Algorithm:

Step 1: Initialize the population size for the candidate solution. For each candidate i the solution is represented as X_i.

Step 2: Objective function f(X) is defined to quantify the X_i.

Step-3: Candidate solutions are updated using the improvement phase and exploration phase as follows:

Improvement Phase

$$X_n = X_0 + r_1.(X_b - X_0) - r_2.(X_w - X_0), \quad (12)$$

Exploration Phase

$$X_n = X_0 + r_3.\left(X_{upper} - X_{lower}\right), \quad (13)$$

where the X_n the new candidate solution X_0 is the old candidate solution, X_b is the best candidate solution found in the population, X_w is the worst solution found so far in the population. X_{upper} is upper bound and X_{lower} is lower bound in the search space. In addition, r_1, r_2, and r_3 are the random vector in the range of [0,1].

Step-4: Find the f(X_n) and if f(X_n)< f(X_0), then update the old solution X_0 by X_n.

Step-5: The maximum iteration T_{max} is checked for obtaining a satisfactory candidate solution.

4 Proposed Approach

The current work employs two different feature selection phases including PSO-GWO and Jaya-GWO. Both phases are then validated with the validation set. To the validated model the set of ML classifiers is then applied to predict the Cancer. Finally, the proposed model is evaluated over 4 different parameters including Accuracy, Precision, Recall, and F-1 Score. Figure 2.shows the proposed model. The steps of working of the proposed model is explained as follows:

Step-1: Consider the dataset for pre-processing using the Standard Scalar method.

Step 2: The dataset is divided into 3 parts Train Data, Test Data, and Validation Set with proportion (70:15:15).

Step 3: Initiate the feature selection phase.

I. Phase-1

a. Apply PSO to the dataset
   i. Initiate the population, Maximum Iteration Tmax
   ii. Calculate the fitness function using the k-fold cross-validation method as follows with VCA as the validation accuracy with K as the number of folds.

$$f() = \frac{\sum_{i=1}^{K} VCA}{K}$$

   iii. Update the particle position using equation 1.
   iv. Revaluate the fitness function f().
   v. If $f_{new}() < f_{old}()$, then keep $f_{new}()$ select the feature, otherwise keep updating the particle position using equation 1.
   vi. Iterate the process for selecting the global best till t<Tmax, where the Tmax is the maximum Feature selection.

b. Apply GWO to the selected features by PSO:
   i. Initiate the Wolf population (α, β, δ, ω), Dimension of search space, Tmax.
   ii. Calculate the fitness function using the k-fold cross-validation method as follows with VCA as the validation accuracy with K as the number of folds.

$$f() = \frac{\sum_{i=1}^{K} VCA}{K}$$

   iii. Initiate encircling the prey phase and update the position of the wolf by using equation 2.
   iv. Start the Hunting phase.
   v. Update the position of α, β, δ wolf using equations 6,7, and 8.
   vi. Update the f() and iterate the process until t<Tmax.

II. Phase II

a. Apply Jaya feature selection algorithm:
   i. Initiate the population size, Tmax, mutation rate, and crossover rate.
   ii. Calculate the fitness function using the k-fold cross-validation method as follows with VCA as the validation accuracy with K as the number of folds.

Table 2. Dataset Description

Name of Dataset	Features	Samples	Classes
ALL-AML	5330	72	ALL
			AML
Ovarian	15155	253	Cancer
			Normal
CNS	6035	102	Tumor
			Normal
Breast	24482	97	Cancer

$$f() = \frac{\sum_{i=1}^{K} VCA}{K}$$

iii. Update the candidate solution using equations 12 and 13.

iv. Iterate the process till t< Tmax.

b. Apply GWO to the selected features by Jaya:

i. Initiate the Wolf population (α, β, δ, ω), Dimension of search space, Tmax.

ii. Calculate the fitness function using the k-fold cross-validation method as follows with VCA as the validation accuracy with K as the number of folds:

$$\mathrm{f}() = \frac{\sum_{\mathrm{i=1}}^{\mathrm{K}} \mathrm{VCA}}{\mathrm{K}}.$$

i. Start encircling the prey phase and revise the position of the wolf by using equation 2.

ii. Start the Hunting phase.

iii. Update the position of α, β, δ wolf utilizing equations 6,7, and 8.

iv. Update the f() and iterate the process until t<Tmax.

Step-4: Apply the ML classifiers to develop a trained model-1 and 2.

4.1 Dataset Description

Different microarray datasets are used in order to measure performance of the proposed model. Table 2 displays the number of dimensions or features and samples or instances with several classes in the datasets such as ALL-AML, Ovarian and Breast cancer dataset, which are considered in [41]. CNS dataset is collected from an internal dataset, which is not publicly available.

This section is devoted to evaluating the performance of the above approach. To start with, the experimental setting is explained and then the results are exposed. The performance of the prosed model is compared with individual meta-heuristic based wrapper algorithms. This work introduces two hybrid feature selection algorithms, namely PSO and GWO 2) The Jaya and GWO algorithms are used for optimal feature selection.

4.2 Classifiers

a) Support Vector Machine

SVMs are reliable supervised machine learning algorithms are suitable for classification as well as regression tasks [4, 31]. The objective is to identify the optimal hyperplane for dividing data points into distinct classes and maximizing the distance between them. SVM may be used to classify linear and non-linear data. Input data is converted into higher-dimensional spaces by using kernel functions.

b) Decision Tree

DT is a flexible and easy-to-understand technique used by many classification and regression applications. The technique uses a recursive algorithm that divides the data into nodes according to its characteristics, resulting in a tree [17]. An internal node in the tree signifies a decision made based on a certain attribute, whereas a leaf node represents a class label or regression value. The goal of doing splits is to augment the process of acquiring information for

Table 3(a). Reduced feature of ALL-AML Dataset

Name of dataset	Features	Algorithm	Feature selected
ALL-AML	5330	PSO	2503
		GA	2211
		GWO	4348
		JAYA	2430
		PSO+GWO	634
		JAYA+GWO	602

Table:3(b). Reduced feature of Ovarian Dataset

Name of dataset	Features	Algorithm	Feature selected
Ovarian	15155	PSO	7100
		GA	6731
		GWO	11377
		JAYA	7063
		PSO+GWO	3779
		JAYA+GWO	2150

Table:3(c). Reduced feature of CNS Dataset

Name of dataset	Features	Algorithm	Feature selected
CNS	6035	PSO	3732
		GA	2580
		GWO	4500
		JAYA	2744
		PSO+GWO	1504
		JAYA+GWO	864

Table:3(d). Reduced feature of Breast Cancer Dataset

Name of dataset	Features	Algorithm	Feature selected
Breast	24482	PSO	7898
		GA	11210
		GWO	19681
		JAYA	11748
		PSO+GWO	6099
		JAYA+GWO	3881

classification or reducing variance for regression at each node. Consequently, a hierarchical model is created, which may accurately represent intricate connections and decision boundaries in the data.

c) Random Forest

RF is an ensemble learning method that improves forecast accuracy and robustness by aggregating the predictions of several decision trees. A series of decision trees are trained on random subsets of the training data and features through bagging and feature randomization techniques. Every tree in the forest independently gathers information about a particular segment of the data and then contributes to the final prediction using a majority vote mechanism for classification tasks.

The Random Forest approach mitigates the problem of overfitting by combining the predictions of several trees, leading to less variability and sustained low bias. The use of this method has established Random Forest as a commonly utilized

Table 4. Result Analysis of Hybrid model with PSO-GWO feature selection algorithm

Dataset	Hybrid Model with PSO+ GWO	ACC	PRE	REC	F1-S
ALL-AML	SVM	91.67	91.89	91.89	91.89
	DT	93.06	92.86	95.12	93.98
	RF	**94.44**	94.29	94.29	94.29
	NB	90.28	91.49	93.48	92.47
	LDA	90.32	85.71	92.31	88.89
Ovarian	SVM	90.91	92.49	94.12	93.29
	DT	90.12	92.18	93.75	92.96
	RF	**92.49**	94.67	94.12	94.40
	NB	91.70	93.83	93.25	93.54
	LDA	89.33	94.12	90.40	92.22
CNS	SVM	**93.14**	94.12	95.52	94.81
	DT	91.18	92.31	93.75	93.02
	RF	91.18	93.94	92.54	93.23
	NB	90.20	93.85	91.04	92.42
	LDA	92.16	95.45	92.65	94.03
Breast	SVM	90.72	92.73	91.07	91.89
	DT	91.75	96.67	90.63	93.55
	RF	93.81	96.49	93.22	94.83
	NB	92.78	95.16	93.65	94.4
	LDA	90.2	92.75	92.75	92.75

Table 5. Result Analysis of Hybrid model with Jaya-GWO feature selection algorithm

Dataset	Hybrid Model with Jaya+ GWO	ACC	PRE	REC	F1-S
ALL-AML	SVM	95.83	94.87	97.37	96.10
	DT	94.44	95.24	95.24	95.24
	RF	**98.22**	98.31	97.39	98.57
	NB	95.83	97.96	96.00	96.97
	LDA	93.06	92.11	94.59	93.33
Ovarian	SVM	**99.81**	98.91	99.45	99.18
	DT	97.62	98.39	98.39	98.39
	RF	98.21	99.44	99.44	99.44
	NB	97.23	98.22	97.65	97.94
	LDA	96.44	97.80	97.27	97.53
CNS	SVM	98.04	98.57	98.57	98.57
	DT	97.06	97.06	98.51	97.78
	RF	**99.0**	98.59	100.00	99.29
	NB	98.04	98.51	98.51	98.51
	LDA	96.08	95.71	98.53	97.10
Breast	SVM	95.88	96.55	96.55	96.55
	DT	94.85	95.31	96.83	96.06
	RF	97.94	98.33	98.33	98.33
	NB	96.91	98.44	96.92	97.67
	LDA	96.81	96.92	98.44	97.67

and highly efficient solution for a range of machine learning challenges [34, 42].

d) Naive Bayes

NB classifiers are fundamental and effective probabilistic machine learning techniques for classification. Bayes' theorem asserts that attributes are statistically independent, hence each attribute influences the probability of the class label. Naive Bayes is very useful in practical situations, especially when dealing with large datasets and where the assumption of independence holds true.

The method uses input features to evaluate the probability of each category and choses the category with the highest likelihood as the predicted label. Naive Bayes stands out due to its computational efficiency, simplicity, and capability to handle both categorical and numerical data [29,42].

e) Latent Dirichlet Allocation

The LDA classifier decreases the number of dimensions in the data and classifies it. Discovering a linear combination of qualities that enhances the diversity between different classes of data while minimizing the diversity within each class may effectively distinguish data classes.

LDA presupposes that the data follows a normal distribution and that each class has a single covariance matrix. Latent Dirichlet Allocation (LDA) is a technique that maps data into a subspace with fewer dimensions.

This process preserves important information about the classes, by which computational complexity can be reduced as well as the accuracy of the classifier can be enhanced. This is very beneficial for data that has several dimensions. It demonstrates exceptional performance in both binary and multiclass classification tasks and is particularly effective at handling scenarios involving many categories [27].

4.3 Validation

The K-fold cross-validation technique is used to test the ability of generalization of model to new data sets. To train the classifiers on each subset, k equal-sized and mutually exclusive subgroups using random splitting are used.

Among the k-datasets, one subset is utilized as the test dataset and the rest subsets are used for training in each iteration. Based on these k tests, we compute accuracy k times, averaging the outcomes to forecast the categorization accuracy.

Cross-validation enhances the reliability of the results by ensuring that every observation in the available data, which are mutually exclusive is utilized for testing. For our study, we partitioned the data into 10 parts of equal size using k = 10. The mean of the outcomes from the 10 iterations was subsequently computed.

5 Empirical Analysis

The proposed model is evaluated in a system with 16 GB of RAB, 500 GB of Hard Disk, 256 GB of SSD, and an Intel i3 processor with Windows 11 OS. The proposed model is evaluated in 2 different phases.

In the first phase, Feature selection algorithms PSO and GWO algorithms are combined and performance is measured. Table 4 shows the result analysis of phase-1. In the next phase Jaya and GWO are combined and performance is measured.

Table 5 shows the result analysis of this phase. To evaluate the effectiveness of our hybrid approach, four microarray gene expression datasets are used in the experiment, with a small number of instances varying between 72 and 253 observations and a large number of features varying between 5000 and 15000 genes.

On one hand, we select datasets of binary type to distinguish healthy patients from cancerous ones, namely leukaemia, central nervous system (CNS), and ovarian containing different types of cancer.

The studied datasets are first divided randomly into two parts: the first 70% of the dataset was used in the first phase of the algorithm to provide weight to each classifier, and the remaining 30% was allocated to test the performance.

The brief explanations of the studied datasets are shown in Table 2.

Table 6. Comparison of Proposed Methods with different state-of-art optimization algorithms

Dataset	Algorithm	Hybrid Model with PSO+ GWO	Hybrid Model with Jaya+ GWO	PSO	GWO	Jaya	GA
ALL-AML	SVM	91.67	95.83	94.6	91.6	87.57	87.85
	DT	93.06	94.44	93.05	91.06	90.38	91.61
	RF	94.44	98.22	97.75	93.62	96.76	95.35
	NB	90.28	95.83	94.2	91.42	92.62	90.18
	LDA	90.32	93.06	92.14	90.33	90.45	92.87
Ovarian	SVM	90.91	99.81	98.81	91.6	98.42	98.42
	DT	90.12	97.62	97.24	92.44	96.45	96.23
	RF	92.49	98.21	93.39	93.42	91.54	93.21
	NB	91.7	97.23	90.54	92.21	90.54	91.35
	LDA	89.33	96.44	90.4	90.15	88.42	89.21
CNS	SVM	93.14	98.04	90.27	90.18	90.19	88.18
	DT	91.18	97.06	86.29	88.30	86.27	79.54
	RF	91.18	99	93.12	91.22	61.53	88.23
	NB	90.2	98.04	62.85	61.90	90.27	61.19
	LDA	92.16	96.08	88.19	89.14	90.19	89.18
Breast	SVM	90.72	95.88	86.78	90.98	92.18	90.16
	DT	91.75	94.85	88.77	91.21	91.89	90.79
	RF	93.81	97.94	90.16	90.79	92.39	91.56
	NB	92.78	96.91	89.91	91.01	92.19	90.76
	LDA	90.2	96.81	86.67	89.79	90.81	89.87

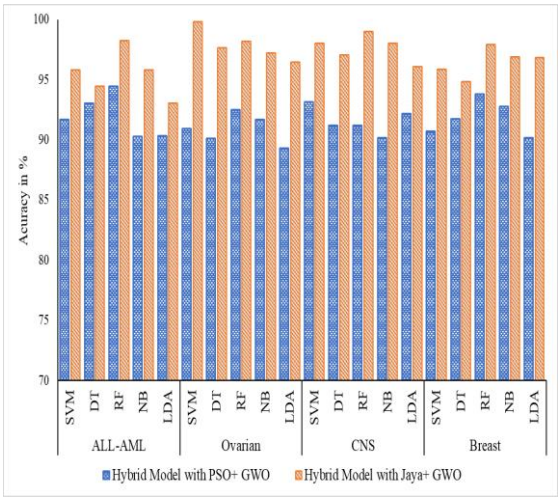

Fig. 4. Accuracy comparison among PSO-GWO and Jaya-GWO

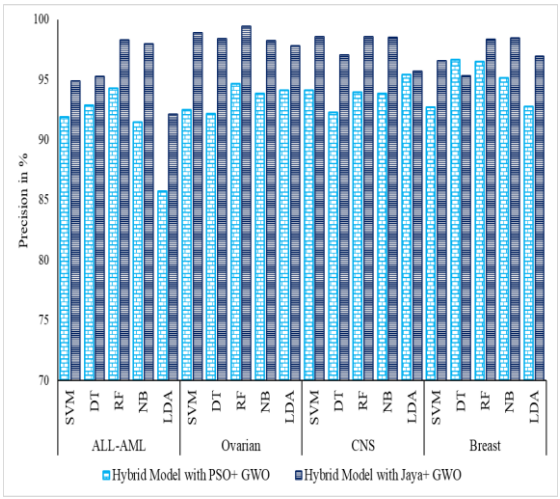

Fig. 5. Precision comparison among PSO-GWO and Jaya-GWO

In terms of classifiers, we selected various types of classifiers to be used in the experiment procedure, namely, SVM, DT, NB, RF and LDA classifiers.

By using a 10-fold cross-validation, the performance of the model was measured. Several performance measures were used to investigate the results of our approach, i.e., accuracy, recall, and precision.

There are values that vary from the highest 100% to the worst 0%. These metrics are defined as follows:

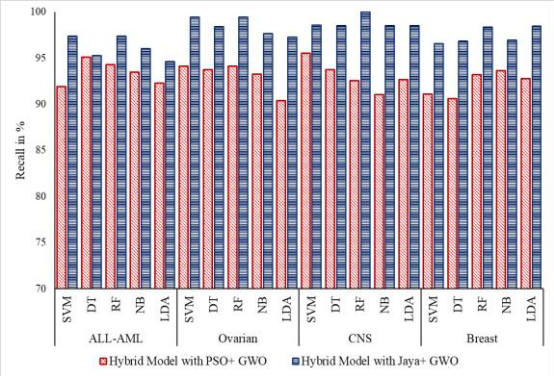

Fig. 6. Recall the comparison between PSO-GWO and Jaya-GWO

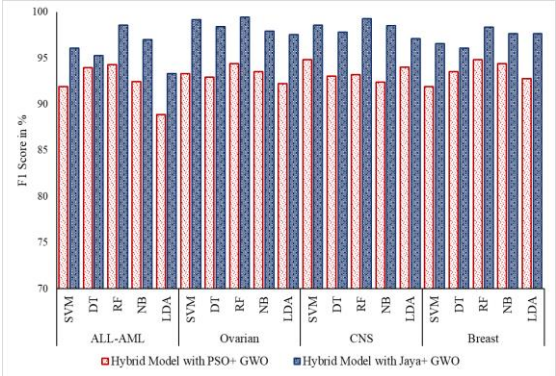

Fig. 7. F1 score the comparison between PSO-GWO and Jaya-GWO

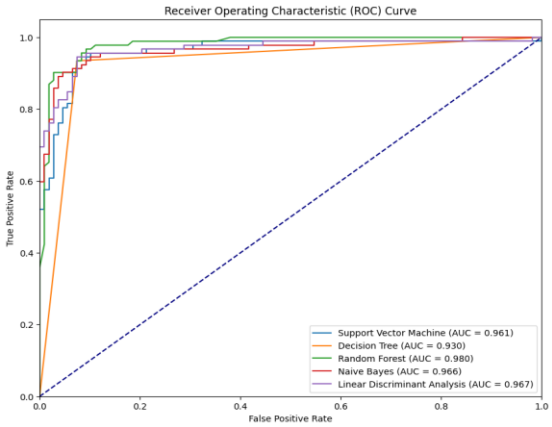

Fig. 8. ROC of proposed Jaya-GWO for ALL-AML dataset

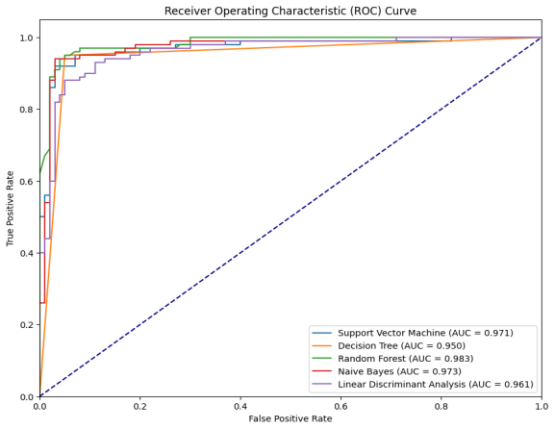

Fig. 9. ROC of proposed Jaya-GWO for Ovarian dataset

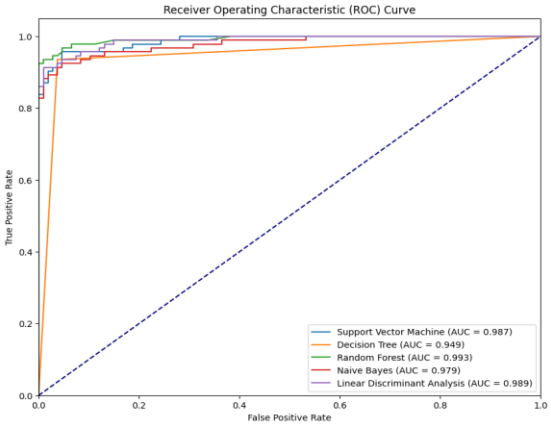

Fig. 10. ROC of proposed Jaya-GWO for CNS dataset

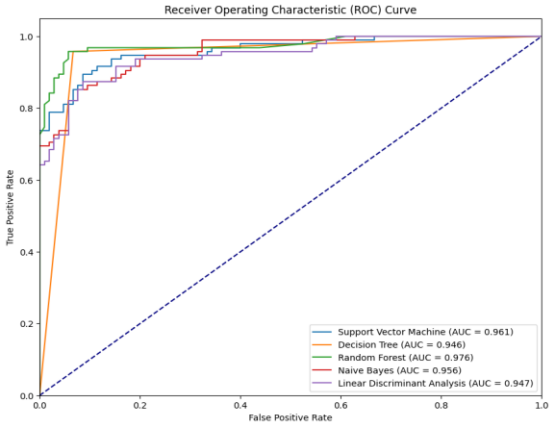

Fig. 11. ROC of proposed Jaya-GWO for Breast cancer dataset

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

$$Recall = \frac{TP}{TP+FP} \times 100,$$

$$Precision = \frac{TP}{TP + FP} \times 100,$$

$$F1\ Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \times 100,$$

As a part of validation phase, we calculate for each dataset the validity measures of each classifier individually, then compare with validity measures of the hybrid approach, which are shown in Tables 4 to 6. Table 3 summarizes the features reduced after applying the wrapper methods to the datasets.

Empirical analysis done earlier indicates that the hybrid model Jaya-GWO performs better than the other models PSO-GWO. Hence to show the efficacy of the proposed Jaya-GWO the ROC analysis is performed with the AUC value. Figures 8 to 11 show the ROC analysis with AUC value for the proposed Jaya-GWO model with different classifiers for ALL-AML, Ovarian, CNS, and Breast Cancer datasets. For the ALL-AML dataset using the proposed Jaya-GWO model the SVM, DT, Rf, NB, and LDA show AUC values of 0.961, 0.93, 0.98, 0.966, and 0.967 respectively. For the Ovarian dataset using the proposed Jaya-GWO model the SVM, DT, RF, NB, and LDA show AUC values of 0.971, 0.95, 0.983, 0.973, and 0.961 respectively. For the CNS dataset using the proposed Jaya-GWO model the SVM, DT, RF, NB, and LDA show an AUC value of 0.987, 0.949, 0.993, 0.979, and 0.989 respectively. Similarly For the Breast Cancer dataset using the proposed Jaya-GWO model the SVM, DT, RF, NB, and LDA show AUC values of 0.961, 0.946, 0.976, 0.956, and 0.947 respectively.

6 Conclusion and Future Research Direction

In this work, a model is built to predict cancer disease by utilizing gene expression microarray data. The challenge of microarray data analysis is the limited number of samples with high-dimensional data, for which there is a need for a reduction of dimensions or features in the original data set. The proposed method involves hybrid wrapper-based methods to generate feature subsets.

Two hybrid approaches PSO-GWO and Jaya-GWO are used to find the optimal feature set from original microarray data. Then five classifiers such as SVM, DT, RF, NB and LDA are used. Our experimental analysis shows that the proposed model Jaya-GWO results in the highest accuracy in comparison to PS0-GWO. The accuracy of the ALL-AML dataset is 98.22%, Accuracy measure is 99.81% for Ovarian, 99% for the CNS dataset and 97.94 for the Breast cancer dataset respectively.

The current work focuses on models used for binary classification. In the future, our objective is to expand the model for multi-class classification by using an ensemble learning model.

References

1. **Hira, Z. M., Gillies, D.F. (2015).** A review of feature selection and feature extraction methods applied on microarray data. Adv. Bioinf., Vol. 2015, No. 198363. DOI: 10.1155/2015/198363.
2. **Ang, J.C., Mirzal, A., Haron, H., Hamed, H.N.A. (2016).** Supervised, unsupervised, and semi-supervised feature selection: A review on gene selection. IEEE/ACM Trans Comput Biol Bioinform, Vol.13, No. 5, pp. 971–989.
3. **Veerabhadrappa-Rangarajan, L. (2010).** Bi-level dimensionality reduction methods using feature selection and feature extraction. International Journal of Computer Applications, Vol. 4, No. 2, pp.33–38.
4. **Bennet, J., Ganaprakasam, C., Kumar, N. (2015).** A hybrid approach for gene selection and classification using support vector machine. International Arab Journal of Information Technology, Vol. 12, pp. 695–700.
5. **Liu, H., Motoda, H. (2007).** Computational methods of feature selection. CRC Press, Boca Raton.
6. **Mitra, P., Murthy, CA., Pal, SK. (2002).** Unsupervised feature selection using feature

similarity. IEEE Trans Pattern Anal Mach Intell, Vol. 24, pp. 301–312.

7. **Shang, W., Huang, H., Zhu, H. (2007).** A novel feature selection algorithm for text categorization. Expert Syst Appl, Vol. 33, pp.1–5.
8. **Yang, J., Honavar, V. (1998).** Feature subset selection using a genetic algorithm. IEEE Intell Syst Appl, Vol. 13, pp. 44–49.
9. **Moradi, P., Gholampour, M. (2016).** A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy. Appl Soft Comput J., Vol. 43, pp. 117–130. DOI: 10.1016/j.asoc.2016.01.044.
10. **Forsati, R., Moayedikia, A., Jensen, R., et al. (2014).** Enriched ant colony optimization and its application in feature selection. Neurocomputing, Vol. 142, pp. 354–371. DOI: 10.1016/j.neucom.2014.03.053.
11. **Gandomi, A.H., Yang, X-S., Alavi, A.H. (2013).** Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. Eng Comput, Vol. 29, pp.17–35.
12. **Karaboga, D., Akay, B. (2009).** A comparative study of artificial bee colony algorithm. Appl Math Comput, Vol. 214, pp. 108–132.
13. **Forsati, R., Moayedikia, A., Keikha, A., Shamsfard, M. (2012).** A novel approach for feature selection based on the bee colony optimization. Int J Comput Appl., Vol. 43, pp. 30–34.
14. **Shukla, A.K., Singh, P., Vardhan, M. (2019).** A new hybrid wrapper TLBO and SA with SVM approach for gene expression datasets. Inf Sci, Vol. 503, pp. 238–254.
15. **Hambali, M.A., Oladele, T.O., Adewole, K. S. (2020).** Microarray cancer feature selection: Review, challenges and research directions. Int. J. Cogn. Comput. Eng., Vol 1, pp. 78–97. DOI: 10.1016/j.ijcce.2020.11.001.
16. **Ghosh, M., Guha, R., Sarkar, R., Abraham, A. (2020).** A wrapper-filter feature selection technique based on ant colony optimization. Neural Comput. Appl., Vol. 32, No. 12, pp. 7839–7857. DOI: 10.1007/s00521-019-04171-3.
17. **Alirezanejad, M., Enayatifar, R., Motameni, H., Nematzadeh, H. (2020).** Heuristic filter feature selection methods for medical datasets. Genomics, Vol. 112, No. 2, pp. 1173–1181. DOI: 10.1016/j.ygeno.2019.07.002.
18. **Goudos, S.K., Yioultsis, T.V., Boursianis, A.D., Psannis, K.E., Siakavara, K. (2019).** Application of New hybrid jaya grey Wolf optimizer to antenna design for 5G communications systems. IEEE Access, Vol. 7, pp. 71061–71071. DOI: 10.1109/ACCESS.2019.2919116.
19. **Abdo, A., Mostafa, R., Abdel-Hamid, L. (2024).** An Optimized Hybrid Approach for Feature Selection Based on Chi-Square and Particle Swarm Optimization Algorithms. Data, Vol. 9, No 2, pp. 20. DOI: 10.3390/data9020020.
20. **Dankolo, M.N., Mohamed-Radzi, N.H., Salehuddin, R., Mustaffa, N.H. (2018).** Hybrid Flower Pollination Algorithm and Support Vector Machine for Breast Cancer Classification. J. Technol. Manag. Bus., Vol. 5, pp. 1, DOI: 10.30880/jtmb.2018.05.01.006.
21. **Bhattacharyya, T., Chatterjee, B., Singh, P. K., Yoon, J.H., Geem, Z.W., Sarkar, R. (2020).** Mayfly in Harmony: A new hybrid meta-heuristic feature selection algorithm. IEEE Access, Vol. 8, pp. 195929–195945. DOI: 10.1109/ACCESS.2020.3031718.
22. **Meenachi, L., Ramakrishnan, en S. (2021).** Metaheuristic Search Based Feature Selection Methods for Classification of Cancer. Pattern Recognit., Vol. 119, pp. 108079. DOI: 10.1016/ j.patcog.2021.108079.
23. **Hameed, S.S., Hassan, W.H., Latiff, L.A., Muhammadsharif, en F.F. (2021).** A comparative study of nature-inspired metaheuristic algorithms using a three-phase hybrid approach for gene selection and classification in high-dimensional cancer datasets. Soft Comput., Vol. 25, No. 13, pp. 8683–8701. DOI: 10.1007/s00500-021-05726-0.

24. Yu, L., Liu, H. (2004). Efficient feature selection via analysis of relevance and redundancy. J Mach Learn Res, Vol. 5, pp. 1205–1224.

25. Alomari Ahmad, O., et al. (2021). Gene selection for microarray data classification based on Gray Wolf Optimizer enhanced with TRIZ-inspired operators. Knowledge-Based Systems, Vol. 223, pp. 107034.

26. Chaudhuri, A., Sahu, T.P. (2021). A hybrid feature selection method based on Binary Jaya algorithm for microarray data classification. Computers & Electrical Engineering, Vol. 90 pp. 106963.

27. Huerta, E.B., Duval, B., Hao, J. (2010). A hybrid LDA and genetic algorithm for gene selection and classification of microarray data. Neurocomputing, Vol. 73, No. 13–15, pp. 2375–383. DOI: 10.1016/j.neucom.2010.03.024.

28. Shukla, A.K., Singh, P., Vardhan, M. (2020). Gene selection for cancer types classification using novel hybrid metaheuristics approach. Swarm and Evolutionary Computation, Vol. 54 pp. 100661.

29. Aziz, Musheer, R. (2022). Nature-inspired metaheuristics model for gene selection and classification of biomedical microarray data. Medical & Biological Engineering & Computing, Vol. 60, No. 6, pp. 1627–1646.

30. Medjahed, S.A., Boukhatem, F. (2024). On the Performance Assessment and Comparison of Features Selection Approaches. Computación y Sistemas, Vol. 28, No. 2, pp. 607–622. DOI: 10.13053/CyS-28-2-4211.

31. Medjahed, S.A., Boukhatem, F. (2024). Applying Support Vector Machines with Different Kernel to Breast Cancer Diagnosis. Computación y Sistemas, Vol. 28, No. 2, pp. 659–667. DOI: 10.13053/CyS-28-2-4207.

32. Pulido, M., Melin, P. (2024). Bird Swarm Algorithm and Particle Swarm Optimization in Ensemble Recurrent Neural Networks Optimization for Time Series Prediction. Computación y Sistemas, Vol. 28, No. 2, pp. 847–859. DOI: 10.13053/CyS-28-2-5027.

33. Nimbiwal, M., Vashishtha, J. (2021). A Novel Hybrid Grey Wolf Optimization Algorithm Using Two-Phase Crossover Approach for Feature Selection and Classification. Computación y Sistemas, Vol. 25, No. 4, pp. 793–801. DOI: 10.13053/CyS-25-4-3931.

34. Safa, B., Mohamed-Hamou, R., Toumouh, A. (2024). Optimizing the Performance of the IDS through Feature-Relevant Selection Using PSO and Random Forest Techniques. Computación y Sistemas, Vol. 28, No. 2, pp. 473–488.

35. Lazar, C., Taminau, J., Meganck, S., Steenhoff, D., Coletta, A., Molter, C., de Schaetzen, V., Duque, R., Bersini, H., Nowe, A. (2012). A survey on filter techniques for feature selection in gene expression microarray analysis. IEEE/ACM Trans. Comput. Biol. Bioinf., Vol. 9, No. 4, pp. 1106–1119.

36. Jadhav, S., He, H., Jenkins, K. (2018). Information Gain Directed Genetic Algorithm Wrapper Feature selection for Credit Rating. Applied Soft Computing. Vol. 69. DOI: 10.1016/j.asoc.2018.04.033.

37. Pirgazi, J., Alimoradi, M., Abharian, T.E., Olyaee, M.H. (2019). An efficient hybrid filter-wrapper Metaheuristic-based gene selection method for high dimensional datasets. Sci. Rep., Vol. 9, No. 1, pp. 18580.

38. Katoch, S., Chauhan, S.S., Kumar, V. (2020). A review on genetic algorithm: Past, present, and future. Multimedia Tools Appl., Vol. 80, pp. 8091–8126.

39. Almugren, N., Alshamlan, H. (2019). A survey on hybrid feature selection methods in microarray gene expression data for cancer classification. IEEE access, Vol. 7 pp. 78533–78548.

40. Sahu, B., Dash, S. (2024). Optimal feature selection from high-dimensional microarray dataset employing hybrid IG-Jaya model. Current Materials Science: Formerly:

Recent Patents on Materials Science, Vol. 17 No. 1, pp. 21–43.

41. **Zhu, Z., Ong, Y.S., Dash, M. (2007).** Markov blanket-embedded genetic algorithm for gene selection. Pattern Recognition, Vol. 40, No. 11, pp. 3236–3248.
42. **Panda, B., Panigrahi, C.R., Pati, B. (2023).** Patient reviews analysis using machine learning. International Journal of Computational Science and Engineering, Vol. 26, No. 2, pp. 111–117.
43. **Behera, S.R., Pati, B., Parida, S. (2023).** RelFFP: An Ensemble Model based on ReliefF and FPA for Efficient Cancer Classification. International Conference in Advances in Power, Signal, and Information Technology (APSIT), pp. 324–329.
44. **Behera, S.R., Pati, B., Parida, S. (2024).** PBb-LMFO: a levy flight integrated MFO inspired ensemble model for cancer diagnosis. International Journal of Information Technology, pp. 1–10.
45. **Bishi, D.K., Padhi, P., Panigrahi, C.R., Pati, B., Rath, C.C. (2023).** Deep learning techniques for hepatocellular carcinoma diagnosis. Computational Intelligence in Cancer Diagnosis, pp. 317–336, Academic Press.
46. **Behera, S.R., Pati, B., Parida, S., Panigrahi, C.R. (2022).** CFS-FA Enabled Ensemble Model for Cancer Diagnosis. International Conference on Machine Learning, Computer Systems and Security (MLCSS), pp. 159–164.

Article received on 23/09/2024; accepted on 13/01/2025.
**Corresponding author is Bibudhendu Pati.*