[Egyptian Informatics Journal 24 (2023) 241–255](https://doi.org/10.1016/j.eij.2023.03.004)

Contents lists available at [ScienceDirect](http://www.sciencedirect.com/science/journal/11108665)

Egyptian Informatics Journal

journal homepage: [www.sciencedirect.com](http://www.sciencedirect.com/)

Full length article

[](http://crossmark.crossref.org/dialog/?doi=10.1016/j.eij.2023.03.004&domain=pdf)An improvised nature-inspired algorithm enfolded broad learning system for disease classification

Pournamasi Parhi [a](#_bookmark0),[⇑](#_bookmark2), Ranjeeta Bisoi [b](#_bookmark1), Pradipta Kishore Dash [b](#_bookmark1)

a *Department of Computer Science Engineering, Siksha ’O’ Anusandhan Deemed to be University, Bhubaneswar, Odisha, India*

b *Multidisciplinary Research Cell, Siksha ’O’ Anusandhan Deemed to be University, Bhubaneswar, Odisha, India*

a r t i c l e i n f o

*Article history:*

Received 22 August 2022

Revised 9 March 2023

Accepted 23 March 2023

Available online 31 March 2023

*Keywords:*

Genomic data

High dimensionality Notable genes Feature extraction Kernel fisher score Classification

Broad learning system Sine–cosine

Improvised monarch butterfly optimization

a b s t r a c t

Deep analysis of genomic data reveals that many deadly diseases are generated due to genetic mutation. To make the health care system more robust, a machine learning researcher’s prime intention is to clas- sify the genomic data more efficiently within less time. As the genomic data suffers from the malediction of excessive dimensionality, the selection of the notable genes is always a big challenge for the researcher. The selection of prominent genomic key attributes by any nature-inspired learning algorithm always remains a non-deterministic polynomial-time (NP-Hard) problem. Therefore, there is always a scope to apply new algorithms. In this projected work, an improvised sine–cosine hybridized Monarch Butterfly Optimization (SC-MBO) algorithm, is embedded with the Broad Learning System (BLS), which is defined as SC-MBO-BLS, for choosing the most significant genes and classifying the genomic data simultaneously. Initially, Kernel-based Fisher Score (K-FS) is applied to select notable genes. Then, the selected genes further undergoes for execution using the SC-MBO-BLS model. To prove the effectiveness of the suggested model, ten cancerous genomic data are considered. Here, several performance evaluators (i.e., precision, MCC, sensitivity, Kappa, F-score, and specificity) are applied for unbiased comparison. This presented model is compared with SC-MBO wrapped Multilayer Perceptron (SC-MBO-MLP), SC-MBO wrapped Extreme Learning Machine (SC-MBO-ELM), and SC-MBO wrapped Kernel Extreme Learning Machine (SC-MBO- KELM) and yields the highest accuracy in ten datasets such as 100%, 98.4%, 99%, 99.6%, 100, 97.2%, 100%, 100%, 98.6%, 99.5% in Leukemia, Colon tumor, Breast cancer, Ovarian cancer, Lymphoma-3, MLL, ALL-AML-3, SRBCT, ALL-AML-4 and Lung cancer respectively. Further, the existing twenty standard models are taken for comparison with the suggested model. Additionally, to assess the presented model, a statistical method i.e., Analysis of variance (ANOVA) is considered. As per the above quantitative and qualitative estimation, it is deduced that the suggested SC-MBO-BLS approach outclasses other considering models.

© 2023 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Computers and Artificial Intel-

ligence, Cairo University. This is an open access article under the CC BY-NC-ND license ([http://creative-](http://creativecommons.org/licenses/by-nc-nd/4.0/)

[commons.org/licenses/by-nc-nd/4.0/](http://creativecommons.org/licenses/by-nc-nd/4.0/)).

1. Introduction

Spending a lot of time and money on genetic research, it has been disclosed that genetic mutation is the root cause of all diseases [[1]](#_bookmark39). In this era of machine learning, researchers’ main focus is to detect these mutated genes quickly and proficiently. If these genes are detected in their initial stage, then diseases can be treatable easily. To make the health care system more efficient, a robust clas- sification model is required to classify the genetic data accurately with less computational complexity. Now-a-days, with the devel- opment of technology, DNA microarray tools provide the facility to monitor the expression levels of various malignant and healthy

\* Corresponding author.

genes in one experiment. Several genes provide inconsistent and noisy signals throughout the whole samples. Therefore, it is a time taking job to analyze all genes as a whole with a small sample size. As a result, a specific analysis is needed to identify the genes which indicate the patterns of expression interrelated with the disease state. So, an artificial intelligence-based diagnostic system has high significance in this genetic research. In the burgeoning of machine learning in various research areas, researchers take the help of var- ious classification techniques to classify the high dimensional microarray data [[1]](#_bookmark39). Moreover, the gene expression datasets are high dimensional datasets having huge numbers of genes or fea- tures but with very fewer numbers of samples. This creates a great drawback in classifying these datasets. Thus, it is always a big chal- lenge for researchers to select the most significant genes in the high

<https://doi.org/10.1016/j.eij.2023.03.004>

1110-8665/© 2023 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Computers and Artificial Intelligence, Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

dimensional microarray data that can help to reduce the computa- tional cost with classificational complexity [[2]](#_bookmark40). Especially two methods are adopted by the researchers to select the significant genes i.e., feature selection and feature extraction. According to the feature extraction technique, the high dimensional feature set is transformed into a reduced lower dimensional feature set by using various linear and nonlinear methods [[3]](#_bookmark41). In the case of the feature selection approach, a subset of the most significant attri- butes is picked up from the high dimensional microarray data by reducing the irrelevant features which have minimal impact on the performance of the learning model. The drawback of the feature extraction technique is the possibility of losing some useful data due to the total transformation of the actual data. Therefore, in this presented work both feature extraction and feature selection are considered for the betterment of the learning rate.

Additionally, the attribute selection [[4]](#_bookmark42), technique is divided of especially three types i.e., Filter, wrapper, and hybrid approach. According to the filter feature selection approach, each feature of the data set is evaluated by applying a statistic measure, then the subset of the most significant feature is selected. But in the wrapper approach, a classifier is used to select the most vital sub- set of features and here classification accuracy is considered to evaluate the most significant feature subset. Though a learning algorithm is considered to evaluate the best feature subset in the wrapper approach so the effectiveness of the wrapper approach is more than the filter approach but the filter approach is also widely accepted due to its less evaluation cost. In the case of the wrapper feature selection method, the combination of metaheuris- tic and machine learning approach is used to select the global best feature subset.

Here, some enfolded approaches are elaborated like the genetic algorithm (GA) is wrapped with a support vector machine (SVM) to get the optimal feature subset of the microarray dataset [[5]](#_bookmark43).GA is wrapped with an Extreme learning machine (ELM) for cancer data classification [[6]](#_bookmark44). Particle swarm optimization (PSO) is embedded with K-Nearest Neighbor (KNN) to select the most relevant feature subset of cancer biomedical data [[7]](#_bookmark45). Metaheuristic algorithm like Genetic Bee Colony (GBC) and Ant Bee Colony (ABC) is used with an SVM classifier for cancer data feature selection [[8,9]](#_bookmark29).GA is wrapped with a Naïve Bayes classifier for gene selection of diabetes data [[10]](#_bookmark29). Breast First Search (BFS) and the very well-known learning algorithm Artificial Neural Network (ANN) are used to select fea- tures and classify the colon data [[11]](#_bookmark29). The metaheuristic Bat algo- rithm (BA) is embedded with the optimum path forest algorithm (OPF) to select the features of medical data [[12]](#_bookmark29). Cat swarm opti- mization (CSO) is wrapped with a very well-known Kernel extreme learning machine (KELM) for the feature selection of medical data [[13]](#_bookmark29). Adaptive genetic algorithms and mutual information are cou- pled together for the feature selection of biomedical data [[14]](#_bookmark29). Though in the wrapper approach classification accuracy is consid- ered a key factor for feature selection so it has more possibilities of obtaining better performance than the filter approach but it has some disadvantages over the filter approach as complex computa- tion and data overfitting. So, by taking the advantage of both the wrapper and filter approach a new feature selection hybrid approach is formed. According to the hybrid strategy, the set of sig- nificant features is first chosen by using the filter feature selection technique, and then out of these selected features again a set of most significant features are selected by using the wrapper approach. For the last ten years, researchers are using both conven- tional and metaheuristic machine learning algorithms to classify various high-dimensional microarray data. The main driving force behind the researcher’s exploration of numerous innovative strate- gies for better results is high classification performance and less evaluation time. SVM [[15–17]](#_bookmark29), ANN [[18]](#_bookmark30), KNN [[20]](#_bookmark31), Fuzzy Set Theory [[19]](#_bookmark32), multi-Layer perception [[23]](#_bookmark33), Functional Link Neural

Network [[21]](#_bookmark34), Backpropagation [[22]](#_bookmark35), Radial Basis Function Neural Network [[24]](#_bookmark36), and others have all been introduced as a classifier to solve the problem.

The above-discussed classifiers are well-liked for their capacity to execute a variety of classification tasks. Deep learning is cur- rently very famous among researchers for its efficiency in classifi- cation purposes. By increasing the number of layers in a neural network, classification performance is effectively promoted. How- ever, the deep network takes a lot of time due to its intricate deep network. A single-layer feed-forward neural network [SLFNN] is better suited to address the challenges since it can solve regression and classification concerns [[25–30]](#_bookmark37). As a learning algorithm, SLFNN is trained using the traditional gradient descent method [[31,32]](#_bookmark46). However, it has several problems, including overfitting, poor con- vergence, and trapping in local minimums [[33]](#_bookmark47).

As a result, the Random-vector Functional-Link Neural Network (RVFLNN), a non-iterative learning technique, is introduced and shown to improve classification performance [[29] and [34]](#_bookmark48). This technique can also get rid of the disadvantage of deep networks’ immense training time. Therefore, RVFLNN requires relatively less training time. However, it has the flaw of not functioning effec- tively in massive, high-volume data [[35]](#_bookmark49). As a result, a novel approach known as the Broad-learning-system (BLS) is suggested by using the idea of the single-layer feed-forward RVFLNN [[36]](#_bookmark50). To improve classification accuracy, in the BLS concept, the input feature nodes are mapped to extended feature nodes that can form a large network. Though, here the input weights are selected at random. Therefore, the classification performance of the algorithm may be impacted. An optimization technique that improves the algorithm’s performance and optimizes the input weight has been offered by researchers as a solution to this problem. Researchers have suggested several *meta*-heuristic algorithms [[37,38]](#_bookmark51) such as PSO [[39] and [40]](#_bookmark52), GA [[44,45]](#_bookmark52), Ant colony optimization (ACO) [[39,40]](#_bookmark52), Moth flame optimization (MFO) [[42,43]](#_bookmark52) and cuckoo search optimization [[41]](#_bookmark52) for optimizing weight and other parameters which can help improve the algorithm performance.

In this research work, the sine–cosine hybridized Monarch But- terfly Optimization algorithm (SC-MBO) is embedded with BLS (SC- MBO-BLS) for the notable gene selection and also for the classifica- tion of genomic data simultaneously. The basic MBO algorithm belongs to the swarm intelligence category which is based on the collective behaviors of self-organized and decentralized systems. MBO algorithm solves many complex optimization issues due to its efficiency and effectiveness. According to google scholar cita- tion, the MBO algorithm [[46]](#_bookmark52) has been mostly applied in different fields of research. In population-oriented optimization methods, there is no certainty of getting a global optimum outcome in a sin- gle run. When the size of iterations (i.e., no. of optimization steps) and the random solutions increase, the possibility of getting a glo- bal optimum result also increases. Every stochastic population- oriented optimization algorithm has to follow the exploration and exploitation phases [[47]](#_bookmark52). These two phases should be balanced properly while solving an optimization problem. Here, in the basic MBO algorithm, the position updating equations are modified using the sine–cosine function to get the optimum result.

By considering the specificity, sensitivity, F-score, and Matthews-correlation-coefficient (MCC) as the validation test, the superiority of the above suggested (SC-MBO-BLS) model is com- pared with that of other established methods such as SC-MBO wrapped multilayer perceptron (SC-MBO-MLP), SC-MBO wrapped Extreme Learning Machine (SC-MBO-ELM), and SC-MBO wrapped Kernel Extreme Learning Machine (SC-MBO-KELM).

This paper’s primary objective is:

* To use kernel-based Fisher Score (K-FS) in the first stage of fea- ture extraction.
* To provide a trustworthy classification model, specifically SC- MBO -BLS for classifying the highly dimensional gene expres- sion data.
* To use minimum features to get high accuracy %.
* Other benchmark approaches like SC-MBO-MLP, SC-MBO-ELM, and SC-MBO-KELM are taken into consideration for comparison to demonstrate the superiority of the provided model.
* In the end, a statistical analysis i.e., the ANNOVA test, was con-

ducted to determine the dominance of the proposed method over other common approaches.

*Zn* = £*n*(*XWtn* + b*tn* ), *n* = 1 · ··· ·· *p* (1)

In Eq. [(1)](#_bookmark5), £*n* describes the *nth* mapping function, *andWtn and*b*tn* point out the random weight and the bias respectively. *Zp* = *Z*1 · · ······· .*Zp* forms a new set of enhancement nodes i.e.,

*Hl* = n*l*(*ZpWht* + b*hl* ), *l* = 1 · ··· ·· .*q* (2)

*Hq* = [*H*1 · .*Hq*]

The resulted output is estimated as

Y=. *z*1 ······ ..*zp*|n1(*zpwh*1 + b*h*1) · ..n*q*(*zpwhm* + b*hm* ) *wq*

*p*

The rest of the work has the following alignments in terms of appearance. The provided model is analyzed in [section 2](#_bookmark3). The back- ground approaches are covered in [section 3](#_bookmark4). While the presented method is covered in [section 4](#_bookmark15). Then the experimental setup is described in [sections 5 and 6](#_bookmark19). Which include the experimentation

= *z* |*Hq* *w*

*p*

*p*

= *z*1 ···· ····· .*zp*|*H*1 · .*Hq wq*

*p q*

*p*

Here, the *wq* is described by

*p*

# (3)

and result validation sections, respectively. Finally, the conclusion is covered in [section 7](#_bookmark38).

*wq* = [*Zp*|*Hq*]+ *Y* (4)

1. Overall analysis of the presented method

*p p*

L2 norm regularized least square problem.

Then,*Aq* = *Zp*|*Hq* and the connected weight *Wq* is estimated by

*Wq* = *argminWq* : *AqWq* — *Y * 2 + k  *Wq * 2

# (5)

[Fig. 1](#_bookmark6) shows the overall analysis of the presented method. Ini- *p*

*p p p z p z*

tially, normalize the dataset by using min–max normalization, then the 10-fold CV method is applied to separate the datasets into training samples and testing samples. In the first stage, the K-FS

approach reduces the features of the datasets to 70%, then the

In Eq. [(5)](#_bookmark5), the constraint constant is k.

*Wq* = k*I* + *AqAqT* —1*AqT Y* (6)

*p*

*p*

*p*

*p*

SC-MBO-BLS algorithm is applied to find the global optimum fea- ture subset with high classification accuracy.

3. Background approaches

Here, *Aq*+ is considered as the inverse of

*Aqi*.*e*., lim k*I* + *AqAqT* —1*Aq*

*p*

*p*

k

*p*

*p*

*p*

# (7)

In this part, all the supported background approaches are elaborated.

1. *1. Broad learning system (BLS)*

[Fig. 2](#_bookmark7) depicts the existing architecture of BLS [[36]](#_bookmark50). The initial attribute set *X* ∈ *RP*×*Q* is mapped arbitrarily to the next feature nodes.

* 1. *Existing MBO algorithm*

This algorithm is based on two operators as migration operator and butterfly adjusting operator [[46,47]](#_bookmark52). The entire population size, taken in this algorithm has been divided into two equal parts. In population 1, the best fitness value of half of the population is kept and the other half of the population is stored in population

2. The best outcome is considered the global best in each iteration

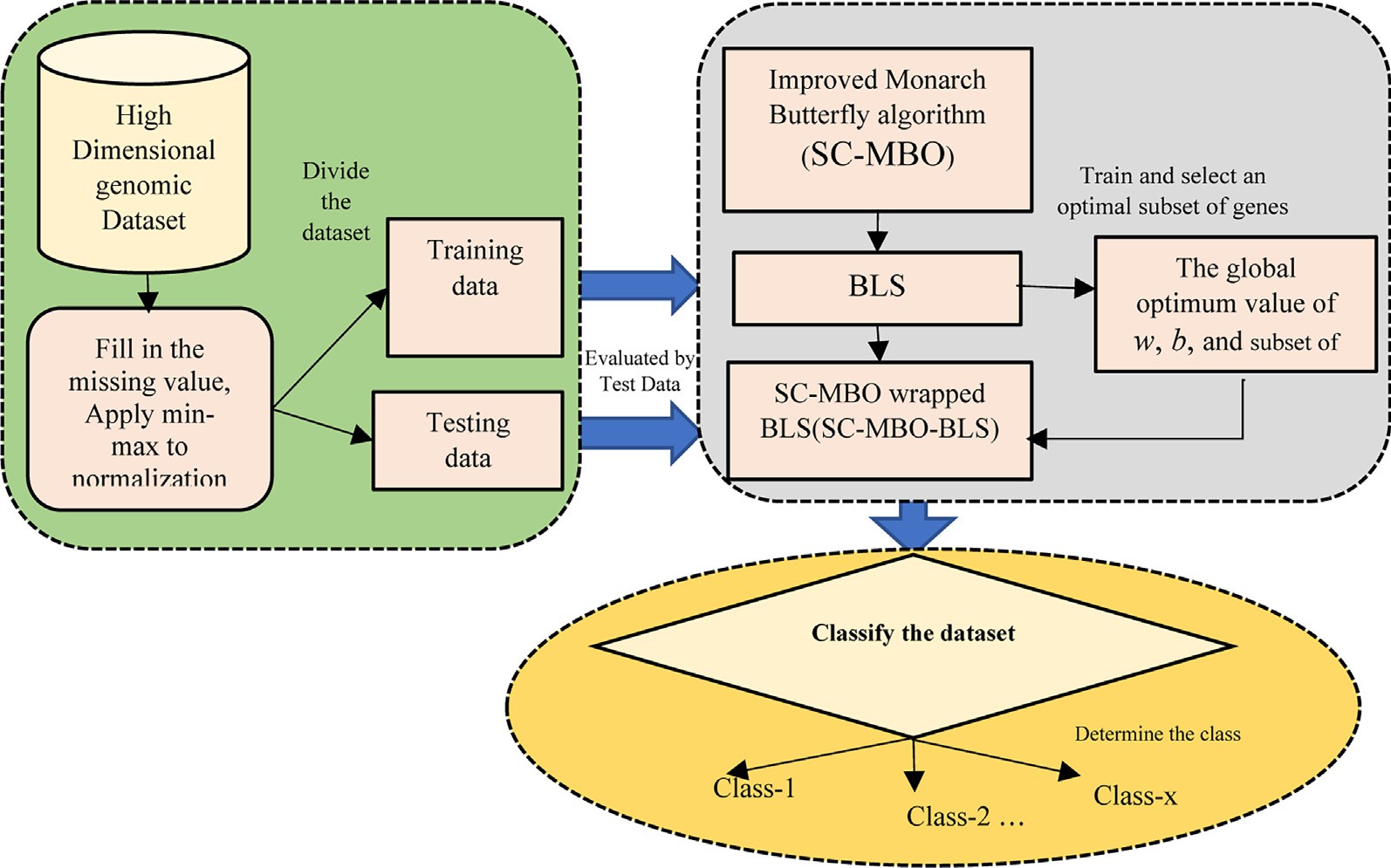


Fig. 1. Basic layout of the presented approach.

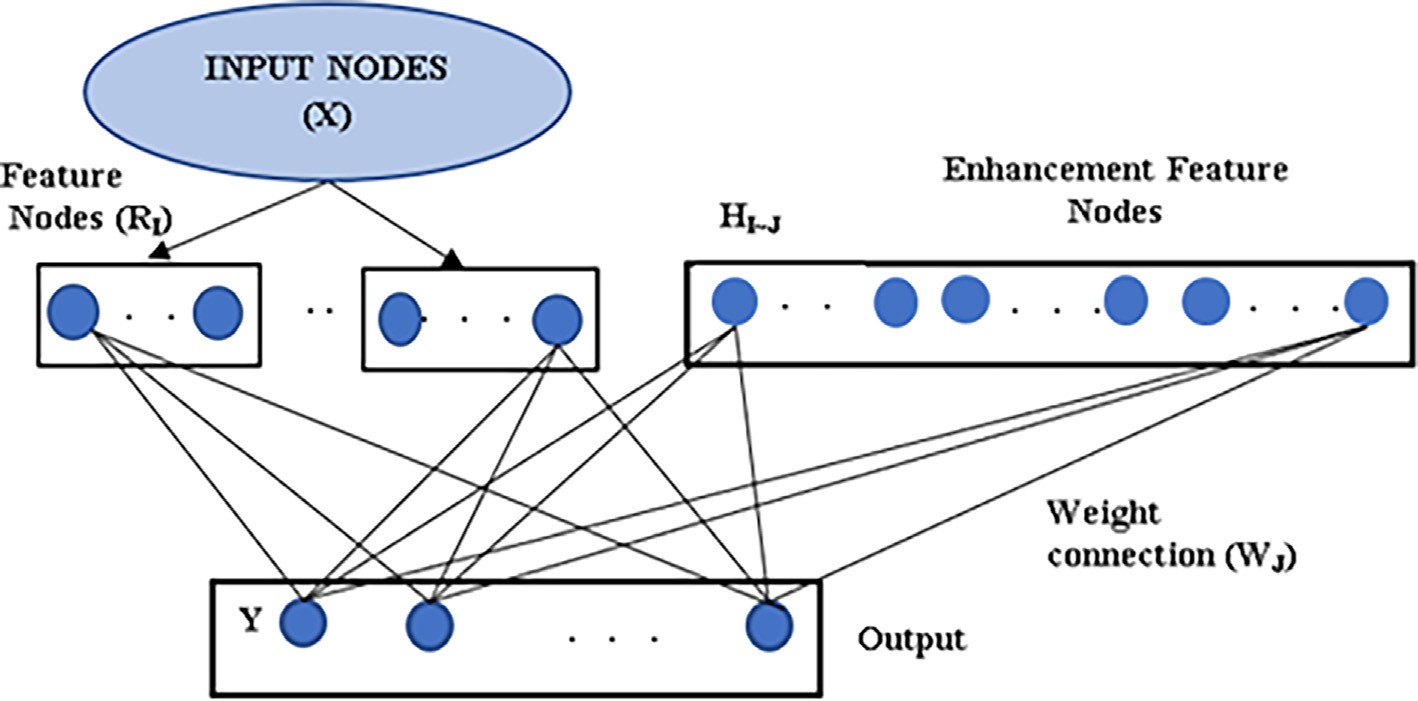


Fig. 2. The basic structure of BLS.

and the updated subpopulation is again mixed with the new pop- ulation. This newly originated population is again split into 2 sub-

*t best*,*n*

where *Y*

eration t, *Yt*

is taken as *nt*h element of the global at the current gen- is the nth element of the arbitrarily chosen butterfly

populations according to the new fitness function and this procedure is continued till the stopping condition is reached. The pictorial view of the MBO is shown in [Fig. 3](#_bookmark11).

* + 1. *Operator-I (Migration Operator)*

By this operator, information is exchanged among both popula- tions. In subpopulation 1, the updation of lth butterfly will be esti- mated as follows:

*i*3,*n*

among subpopulation 2. The weighted factor (n) will be estimated as follows:

n = *Zmax*/*t*2 (10)

Here, *Zmax* = max no. of walk step of each butterfly in every step and *t* shows the current generation.

In Eq. [(11)](#_bookmark8), *dYn* points out the steps of each butterfly and this is estimated by applying Levy flight approach.

*t*+1 *l*

*Y*

=

*Y*

*t*

*i*1,*n if r*<*p*

( *Y*

(8)

*dx* = *Levy* *xt*

# (11)

,*n*

*i*2,*n*

*else n n*

In t + 1 generation, *Yt*+1 points out the location of *Yl* in lth dimension where *l*1 and *l*2 are the integer indexes arbitrarily selected among subpopulations 1 and 2. Here, *r* is considered as the multiplication of random real numbers between (0,1) and migration period.

*l*,*n*

* + 1. *Operator-II (Adjusting Operator)*

The change of position of every butterfly in subpopulation 2 is

* 1. *Sine-Cosine mechanism*

The sine–cosine (SC) algorithm is developed for solving opti- mization issues by Mirjalili [[49]](#_bookmark53). In this algorithm, the positions of the particles are upgraded by applying the sine and cosine mechanisms. This algorithm follows the below two equations (i.e., Eqs. [(12)](#_bookmark9) and [(13)](#_bookmark10)) to update the solutions.

estimated *based on p* (i.e., adjusting ratio) and *BAR* (i.e., the adjust-

*Ax*+1 = *Ax*

+ *a* × sin(*a* )× *a Bx*

— *zx*

# , *a* < 0.5 (12)

ing rate of butterfly)

# 8>

*t best*,

*Y*

## *nifRand* ≤ *p*

*m*,*k*

*Ax*+1

*m*,*k*

*Ax*

*a* × cos(*a* )× *a*

*m*,*k*

*Bx*

*m*,*k*

— *zx*

, *a*

# < 0.5 (13)

*t*+1 *l*,*n*

1 2 3

4

*Y* = <

*Y*

*m*,*k*

*t i*3,*n*

*ifRand* > *p*K*Rand* ≤ *BAR*

(9)

*m*,*k* =

*m*,*k* + 1

2 3 *m*,*k*

*m*,*k* 4

>: *Yt* + n × (*dYn* — 0.5)*ifRand* > *p*K*Rand* > *BAR*

*l*,*n*

Here, *Ax* represents the current position at time ×, in the

whole population, *Bx*

*m*,*k*

represents the location of the best solution,

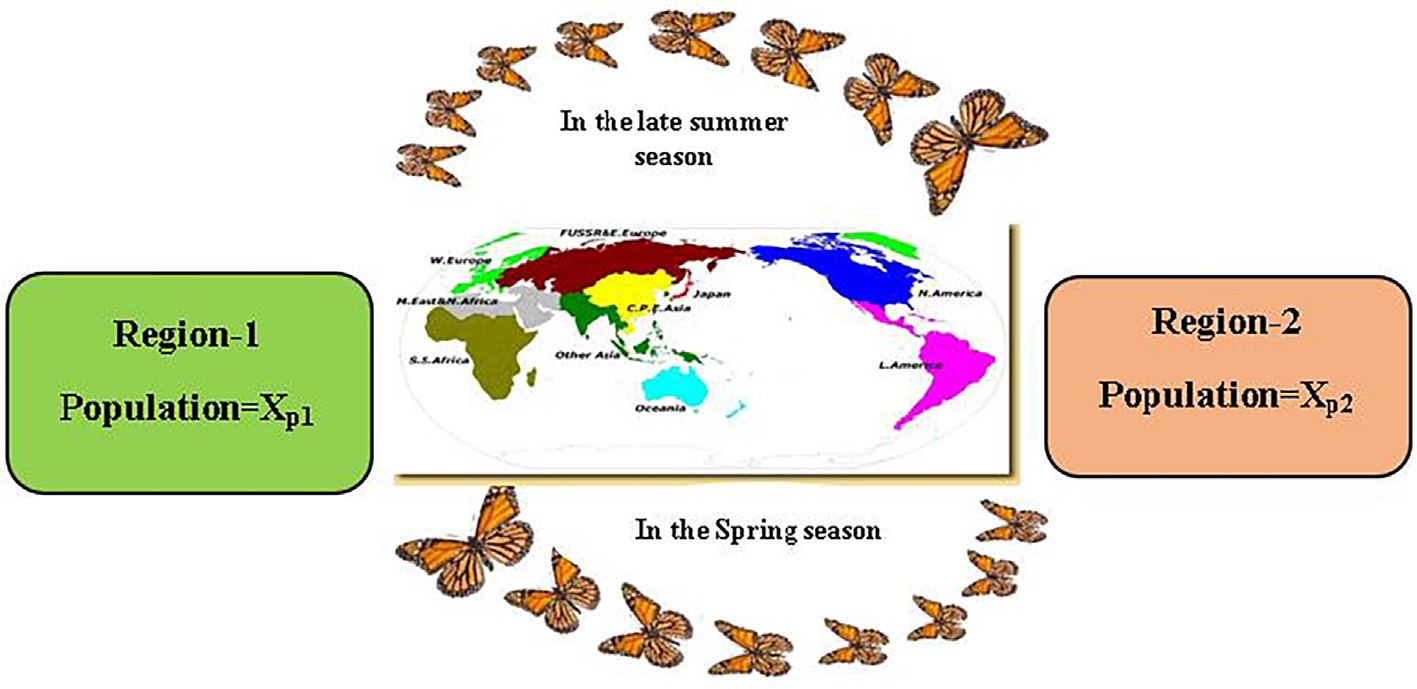


Fig 3. Pictorial view of the Monarch Butterfly.

*a*1*, a*2*,* and *a*3 denote randomly generated numbers but *a*4 is the random number generated between 0 and1.

The solution’s next position will be placed among the current solution, then the desired point can be decided by applying *a*1 (i.e., *a*1 *<* 1, in the *exploitation phase*) or beyond this space (i.e., *a1 >* 1, in the *exploration phase*).

*a*1 = *l*(1 — *x*/*xmax*) (14)

Algorithm 2: Adjusting operator (SC-MBO)

Start

for n = 1 to N2(no. of population (i.e., butterflies in Land2) present)

for k = 1 to j (no. of elements in n*th* monarch butterfly)

if r ≤ p then,

*Ax*+1 = *Ax*

here *rand* ~ *U (01)*(21)

In Eq. [(14)](#_bookmark12), *x* represents the current iteration, *xmax* denotes the

*n*,*z*

*best*,*k*

maximum size of iteration, and in the distance formulation, a ran-

dom weight is associated with the desired point to learn the impact of that point [[49]](#_bookmark53).

elseif rand () < 0.5

*Ax*+1 = *Ax*

*n*,*k*

*n*,*k*

+ *a* × sin(*a* )× *a Bx*

1 2 3

*n*,*k*

— *Ax*

*n*,*k*

(22)

else

*Ax*+1 = *Ax*

+ *a* × sin(*a* )× *a Bx*

— *Ax*

(23)

* 1. *Sine-Cosine hybridized Monarch butterfly optimization (SC-MBO)*

The reason behind the merging of MBO and SCA algorithms is

end if

*n*,*k*

end if

*n*,*k*

*n*,*k*

*n*,*k*

described in this section. The prime benefit of SCA is its extraordi- nary exploration capacity. But in some cases, SCA fails in balancing

1 2 3

if r>p then

*Ax*+1 = *Ax*+1Here, r3 ~ U [1,2, 3,.. ., N2 ](24)

*n*,*k r*3,*k*

between the exploitation and exploration phase which creates sub- elseif rand () < 0.5

1 2 3

*n*,*k*

else

*n*,*k*

*n*,*k*

*n*,*k*

optimal results. In a few cases, SCA skips the global best solution

and small exploitation happened. This mechanism reduces the search capability of the algorithm. Meanwhile, at the time of the

*Ax*+1 = *Ax*

+ *a* × sin(*a* )× *a Bx*

— *zx*

(25)

search process, the MBO algorithm is quite efficient in balancing

*Ax*+1 = *Ax*

+ *a* × cos(*a* )× *a Bx*

— *zx*

(26)

the exploration and exploitation phase. Further, like other evolu-

tionary algorithms, MBO may be trapped at local optima. There- fore, a hybridization of the two algorithms namely the SC-based MBO (i.e., SC-MBO) algorithm is proposed here to resolve the above

*n*,*k*

*n*,*z*

*n*,*k*

end if

if r> R

*n*,*k*

end if

*Ax*+1 = *Ax*

+ b ×((*dck*) — 0.5))(27)

1 2 3

issues. The pseudo-codes SC-MBO are described in algorithm 1 and algorithm 2.

*n*,*k*

end

*n*,*k*

Algorithm 1: Migration operator (SC-MBO)

Start

for m = 1 to N1(no. of population (i.e., butterflies in Land1) present)

for k = 1 to j (no. of elements in *mth* monarch butterfly) r = rand \* S, here, rand ~ U (0,1)

if r ≤ q then

*Ax*+1 = *Ax*+1 Here, r ~ U [1,2,3, .. ., N1] *Ax*+1 = *Ax* (15)

end for k

end for n

End

In the above pseudocode (i.e., Algorithm 2), *Ax*+1 shows the *k*th ele- ment of *An* in × + 1 generation and denotes the n position of MB in Land2. *Ax* shows the *k*th element of *Ar* in *xth* generation and denotes the currently updated position of *r*3 MB. *dck* represents

*n*,*k*

*r*3 ,*k*

3

*m*,*k*

*r*1,*k* 1

elseif rand () < 0.5

1 2 3

*m*,*k*

*m*,*k*

*m*,*k*

*d*1,*k*

the walk step of *nth* MB and will be calculated by applying Levy flight mechanism. Here, b *= IBmax*/*x*, where *IBmax* = no. of steps of

*Ax*+1 = *Ax*

else

*m*,*k*

*x*,*k*

+ *a* × sin(*a* )× *a Bx*

— *Ax*

(16)

each butterfly in single step.

*Ax*+1 = *Ax*

+ *a* × sin(*a* )× *a Bx*

— *Ax*

(17)

* 1. *Kernel-Fisher score*

*m*,*k*

end if end if

if r > q then

*m*,*k*

*m*,*k*

*m*,*k*

In the basic Fisher score (FS) feature extraction method, a fea- ture is selected as per its score which is calculated according to

Eq. [(28)](#_bookmark13). Initially, a M × N input matrix is taken where M point outs

1 2 3

*Ax*+1 = *Ax*+1 Here, r2 ~ U [1,2, 3, .. ., N1] (18)

*m*,*k*

*r*2,*k*

the number of genes and N presents the sample size. Then, in Ker-

elseif rand () < 0.5

1 2 3

nel Fisher score (K-FS) [[48]](#_bookmark52) of each gene is computed and the mean

*Ax*+1 = *Ax*

+ *a* × sin(*a* )× *a Bx*

— *zx*

(19)

value (i.e., a threshold value (THV)) is estimated. The gene with a

else

*m*,*k*

*Ax*+1 = *Ax*

*m*,*k*

+ *a*

× cos(*a* )× *a Bx*

*m*,*k*

— *zx*

*m*,*k*

(20) end if

higher score than THV will be kept in the feature space and the gene with a lower score than THV will be discarded.

*m*,*k*

end if

*m*,*k* 1

2 3 *m*,*k*

*m*,*k*

—(+) — 2

—(—) — 2

end for k

end for m

End

*GS* *Aj* =

*f j* — *f j*

# P 2

1

*n*

*f* (+) — *f* —(+)

(*n*+—1)

*x*=1

*x*,*j*

*j*

+ *f j* — *f j*

# P

— 1

*n*—

*x*=1

(*n*— —1)

# 2

*f* (—) — *f* —(—)

*x*,*j*

*j*

(28)

In the above pseudocode (i.e., Algorithm 1), *Ax*+1 shows the *k*th ele- ment of *Am* at × + 1 generation and denotes the m position of MB

*m*,*k*

—

In the above Eq. [(28)](#_bookmark13), *f x* is taken as the training vector, *f j* is con- sidered as the *jth* attribute of the datasets, *n*+ and *n*— are assumed as

(i.e., monarch butterfly) in Land1. *Ax*

shows the *kth* element of *Ar*

—(+) *th*

*r*1 ,*k*

1 the + ve and -ve instances, *f j* is taken as the *j*

attribute of

in xth generation and denotes the currently changed position of *r*1

2

—(—)

*x r*2 ,*k*

MB. *A*

shows the *k*th element of *Ar*

in xth generation and denotes

the + ve value of the datasets and *f j* is taken as the *jth* attribute

the currently updated position of *r*2 MB. of the -ve value of the datasets. Likewise, *f* (+) is assumed as the

*x*,*j*



**Start**

Yes

**If Fisher score>Threshold (mean value calculated from all Fisher score**

No

End

Remove this feature from the feature space.

Add this feature in the feature space.

Compute FS value of each gene in the dataset

Transfer the Dataset into Kernel space using Kernel function (Linear Kernel/ RBF Kernel)

Load Dataset

*jth* attribute of the *xth* + ve valued instances and —(—)

Fig. 4. Flow chart of KFS.

*f j* is assumed as the *jth* gene or attribute of the *xth*, -ve valued instances.

The mutual information between genes is not taken in standard FS. To avoid this demerit, K-FS [[48]](#_bookmark52) transforms non-linearly separa- ble data into linearly separable data using a kernel function which decreases the computational overhead. The stepwise flow chart of K-FS is depicted in [Fig. 4](#_bookmark14).

1. Suggested methods

*4.1. Proposed SC-MBO-BLS algorithm*

Initially, K-FS [[48,49]](#_bookmark52) approach is used to extract the most sig- nificant genes. In this method, each gene is associated with a rank. This technique selects up to 500 genes [[50]](#_bookmark54) among thousands of genes in the dataset. These preselected genes are forwarded to SC-MBO-BLS for getting the best gene subset.

In this presented approach, the SC-MBO-BLS technique is used to optimize the bias (b) and weight (w) of BLS and to find the opti- mum gene subset simultaneously. [Fig. 5](#_bookmark17) and algorithm 3 give a complete framework of the presented model.

Let us consider, *Xj* = n*w*, *b*, *Xj* , *Xj* , *Xj* , ··· ., *Xj* o as an individual

1

2

3

*K*

Algorithm:3 Suggested SC-MBO-BLS

Input: Set size of population (P), ratio of migration RM, Peri (i.e., migration period), Adjusting Rate of Butterfly (AR), max step of walk of Levy flight *Zmax* and iteration size (IS), Fitness (f), Upper and Lower bound (i.e., UpB and LoB)

Output: Classification accuracy percentage and the length of the subset of genes.

S1: for each independent solution repeat step 2 and 3

S2: By applying logistic function, the value of every gene is converted into 0 and 1. Here, 1 is considered as selection and 0 is considered as rejection of that particular gene.

S3. Find out the fitness by applying w, b, and gene subsets S4: end for

5: Keep aside the fitness values in descending order and select the best value and worst value.

6: According to fitness value, set the population size (P). 7: Identify the location of the best solution.

8: Equate the mean of the fitness values.

9: while X< IS do

10: if X == 1 then

11: for Y=1: P do

12: Update w, b by using Eqs. (15) -(27).

13: end for

14: else if (meanFit\_ curr\_ solution – meanFit\_prev\_solution) / meanFit\_ curr\_ solution > .001

15: Then

solution with K dimensional attribute and j={1, 2, 3, ··· ., *K*}. In this

set of solutions, the first two bits are reserved for *w* and *b* and the other bits are considered as gene subsets. Here, 0 expresses the rejection and 1 expresses the selection of the genomic attribute subset. Therefore, *Xj* is shown as *Xj* = [*w*, *b*, 1, 0, 0, ··· .., 1].

By applying a logistic function, the value of every gene is

expressed in Eq. (29)-(31)

( 1 *logsig* *xq* 0 5

*xq* =

,

*p*

>

.

(29)

Replicate steps 11 to step14

16: Else

17: break.

18: end if

19: end if

20: for each single improved solution do

21: Mark the LoB and UpB of the solution place, bias(b), and weight (w).

22: Repeat steps 1 to 4 for getting the new values of fitness.

23: if fit\_current\_sol > fit\_previous\_sol, then

24: Upgrade the fitness value of the sol and its place, bias(b),

*p* 0, *otherwise*

In Eq. [(29)](#_bookmark16),

and weight (w).

26: Else

27: Set out the prev\_sol\_ fit.

28: Store the solution place, bias(b), and weight (w)..

*logsig* *xq* = 1

*p*

(—*ym*)

# (30)

29: end if

1 + *e k*

P10 *test Acci*

*fit*(*value*) = *Acc*(*avg*) =

*i*=1 10 (31)

31: end for

32: end while

30: Replicate the step 5 to step8

33: Get the concluding result as CA (classification accuracy) %

with the feature subset length.

Begin



Replace the missing cells and apply min-max normalization

Split the whole samples employing a 10-fold CV approach

Training Sample1 Testing Sample1 …… Training Sample10 Testing Sample10

Training Samples

Testing Samples

Extract genes using K-FS

Initialize NP populations, sensory modality (c) value, stimulus intensity (I), and power exponent (a) of the SC- MBO

If a gene subgroup is carrying D-dimension similar to the extracted genes, then every candidate solution is,

*X j* = { }

Here, j = {1, 2, 3 …, M}

If the solution having the size n bit, then the first 2 bits are reserved for *w,*b, and n-2 bits are coded between (1,0) for a subset of gene employing a transformation function.

Here, 1 and 0 specify the selection and rejection of that gene respectively.

*Xi* = [w, b, 0, 1, 1, …., 0]

Provide optimal w, b, and feature subset to BLS in the testing set

Code random parameters (w, b) and the subset of genes

Classify the test data using the SC-MBO-BLS model

If finished 10 runs?

Consider Average accuracy of BLS as fitness function

Find out the fitness value of the individual solution of training sample

No

Observe the best one and worst one in the population

Yes

Sort the fitness values

Find the classification result of each run

Change the solution employing Eqs. (15)-

(27) of the SC-MBO

Yes

No

If reached the maximum no. iterations?

Stop

Store the optimum w, b, and subset of gene

Fig. 5. Pictorial representation of the SC-MBO-BLS model.

Table 1

Explanation of 10 standard datasets.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Datasets | Dimensions | Sample size | No. of Genes | Classes |
| Leukemia [[51]](#_bookmark55) | 72 × 7129 | 72 | 7129 | 2 |
| Colon tumor [[52]](#_bookmark56) | 62 × 2000 | 62 | 2000 | 2 |
| Breast cancer [[53]](#_bookmark62) | 97 × 24482 | 97 | 24,482 | 2 |
| Ovarian cancer [[54]](#_bookmark65) | 253 × 15154 | 253 | 15,154 | 2 |
| MLL [[53]](#_bookmark62) | 72 × 12582 | 72 | 12,582 | 3 |
| Lymphoma-3[[55]](#_bookmark57) | 62 × 4026 | 62 | 4026 | 3 |
| SRBCT [[56]](#_bookmark57) | 88 × 2308 | 88 | 2308 | 4 |
| ALL-AML-3 [[57]](#_bookmark57) | 72 × 7129 | 72 | 7129 | 3 |
| ALL-AML-4 [[53]](#_bookmark62) | 72 × 7129 | 72 | 7129 | 4 |
| Lung cancer [[55]](#_bookmark57) | 203 × 12600 | 203 | 12,600 | 5 |

Table 2

Default Parameter value taken in all the algorithms.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| MLP | ELM | KELM | BLS | Sine-Cosine | MBO |
| No. of Iterations | Default size of | Default size of max | No. of iterations (Max) = 100We | Default size of | Population |
| (Max) = 100Number of HL | population = 100 | iterations = 100 | (coefficients of feature nodes) | population = 100 | size = 100 |
| (Hidden | Default size of max | C and c value = [2–7, | =rand | Default size of max | Default size of max |
| Layers) = 3Number of HN | iterations = 100 | 2–8, .. ., 27, 28] | (0,1) | iterations = 100Constant | iterations = 100 |
| (Hidden Nodes)  in every layer = 5 | w =rand (0,1)  b = rand (0,1) |  | Wh (coefficients of enhancement nodes) = rand (0,1)  b = rand (0,1)  Enhancement Nodes = 100N1 (number of feature nodes per window) = 100N2 | (l) = 3.0  a1 = 20  a2 = rand(0,2p) a3 = 20  a4 = rand (0,1) | Default period of migration = 1.2 Default ratio of migration = 5/12AR (Default Adjusting  rate) = 5/12 |
|  |  |  | (number of windows of feature |  | Default size of max |
|  |  |  | nodes) = 100 |  | generation = 100 |
|  |  |  |  |  | Default size of max |
|  |  |  |  |  | step walk = 1 |

1. Experimental details
   1. *Execution environment*

The whole work is carried out under the following execution environment:

Table 3

K-FS extracted eminent genomic attributes and its accuracy percentage in three binary datasets.

CPU: Intel(R) Core (TM), Processing speed of CPU is 2 GHz, OS used: Windows 10, Language: MATLAB, version- R2020A and Ran- dom Access Memory: 8 GB DDR2 RAM.

* 1. *Detailed description of used dataset*

The detailed elaboration of the used datasets is shown in [Table 1](#_bookmark18).

* 1. *Default parameters*

The default values of the all models parameters are described in [Table 2](#_bookmark18).

* 1. *Model evaluation parameters*

*Sensitivity*(*Sn*)= True Positive (TP)/

# ( True Positive (TP)+ False Negative(FN)) (32)

*Specificity*(*Sp*)=True Nagative(TN) /

# (True Nagative(TN)+ False Positive(*FP*)) (33)

* + - Precision (Pr):

*Pr* = *TP*/(*TP* + *FP*) (34)

* + - * F-Score (Fs):

|  |  |  |
| --- | --- | --- |
| Dataset | # Eminent Genes | ACC percentage |
| Leukemia | 5 | 75.76 |
|  | 10 | 83.19 |
|  | 50 | 94.62 |
|  | 100 | 97.32 |
|  | 200 | 98.82 |
|  | 500 | 98.57 |
| Colon tumor | 5 | 73.62 |
|  | 10 | 78.58 |
|  | 50 | 86.26 |
|  | 100 | 88.7 |
|  | 200 | 89.72 |
|  | 500 | 87.82 |
| Breast Cancer | 5 | 85.62 |
|  | 10 | 90.12 |
|  | 50 | 93 |
|  | 100 | 96.72 |
|  | 200 | 97.62 |
|  | 500 | 96.8 |
| Ovarian cancer | 5 | 82.63 |
|  | 10 | 92.82 |
|  | 50 | 96.75 |
|  | 100 | 98.85 |
|  | 200 | 97.7 |
|  | 500 | 97.85 |

*Fs* = 2 × *Pr* × *Sn*

*Pr* + *Sn*

# (35)

Table 4

K-FS extracted eminent genomic attributes and its accuracy percentage in three multi-class datasets.

Datasets # Eminent Genes ACC percentage

Lymphoma-3 5 80.82

10 96.15

50 96.74

100 98.2

200 100

500 99.82

MLL 5 79.96

10 85.21

## *Obser ed Acc TP* + *TN*

*TP* + *FP* + *TN* + *FN*

*v* =

(*TP*+*FN*)×(*TP*+*FP*) + (*FP*+*TN*)×(*FN*+*TN*)

*Expected Acc* = *TP*+*FP*+*TN*+*FN TP*+*FP*+*TN*+*FN*

*TP* + *FP* + *TN* + *FN*

1. Result discussion

*6.1. Pre-extraction of notable genes*

# (38)

(39)

50 88.9

100 94.77

200 97.85

500 97.65

ALL-AML-3 5 52.86

|  |  |  |
| --- | --- | --- |
| 10 | 51.2 |  |
| 50 | 89.8 |
| 100 | 94.46 | *6.2. Results obtained from SC-MBO-BLS* |
| 200 | 93.78 |  |

Initially, K-FS approach is used to pre-extract the notable genes. The most prominent *N* number of genes (i.e., upto1500 genes) are extracted from the thousand’s genes of the entire dataset. [Tables 3](#_bookmark20) [and 4](#_bookmark20) illustrate topmost extracted genes and its accuracy % of all the datasets.

By using K-FS pre-extraction approach, 100 prominent attri- butes of Ovarian, 200 prominent attributes of Leukemia, colon tumor, breast cancer, Lymphoma, MLL, and SRBCT, 500 prominent attributes of ALL-MLL3, ALL-MLL4 and 1000 prominent attributes of Lung cancer are extracted. These extracted genes are then for- warded to SC-MBO-BLS model.

|  |  |  |
| --- | --- | --- |
|  | 500 | 94.86 |
| SRBCT | 5 | 59.97 |
|  | 10 | 66.83 |
|  | 50 | 89.53 |
|  | 100 | 98.48 |
|  | 200 | 100 |
|  | 500 | 100 |
| ALL-AML-4 | 5 | 64.83 |
|  | 10 | 68.36 |
|  | 50 | 88.98 |
|  | 100 | 92.64 |
|  | 200 | 93.82 |
|  | 500 | 94.9 |
| Lung cancer | 5 | 86.56 |
|  | 10 | 91.5 |
|  | 50 | 92.45 |
|  | 100 | 94.3 |
|  | 200 | 95.07 |
|  | 500 | 95.62 |

In [Table 5](#_bookmark21), all performance evaluators of ten microarray data- sets like acc%, sensitivity, precision, F-Score, MCC, specificity, and Kappa are noted. According to [Table 5](#_bookmark21), it is observed that Leuke- mia, Lymphoma-3, SRBCT, and ALL-AML-3 datasets outperform the others by resulting 100% accuracy. Here, negative samples of Leukemia, Colon, Breast cancer, and Lymphoma-3 are very well classified by giving a specificity rate 100%. Similarly, positive sam- ples of Leukemia, Ovarian, SRBCT, and ALL-AML-4 are very well classified by giving a sensitivity rate 100%. The confusion matrix of all ten datasets is illustrated in [Fig. 6](#_bookmark22)(a)-4(j).

|  |  |  |
| --- | --- | --- |
| 1000 | 95.86 | Here, the convergence graph of SC-MBO-MLP, SC-MBO-ELM, SC- |
| 1200 | 95.2 | MBO-KELM, and SC-MBO-BLS approaches in ten-microarray data- |

1500 95

* Matthews correlation coefficient (MCC):

sets are depicted in [Fig. 7](#_bookmark23)(a)-4(j). It is noted that the accuracy per- centages of ten datasets are incremented continuously up to 100 iterations. In the case of Leukemia, the accuracy is converging after

the 50*th*, 55*th*, 84*th*, and 89*th* iterations in SC-MBO-BLS, SC-MBO-

*MCC*

(*TP* × *TN*) — (*FP* × *FN*)

KELM, SC-MBO-ELM, and SC-MBO-MLP models respectively. In

# 36

= p(ﬃﬃ*T*ﬃﬃﬃ*P*ﬃﬃﬃﬃ+ﬃﬃﬃﬃ*F*ﬃﬃﬃ*P*ﬃﬃﬃ)ﬃ(*TP* + *FN*)(*TN* + *FP*)(*TN* + *FN*) ( )

* Kappa (Kpa):

## (*Observed Acc* — *Expected Acc*)

Colon, the accuracy is converging after 44*th*, 66*th*, 78*th*, and 84*th* iter- ations in SC-MBO-BLS, SC-MBO-KELM, SC-MBO-ELM, and SC-MBO-

MLP models respectively. In Breast cancer data, the accuracy is

converged after 44*th*, 65*th*, 73*th*, and 80*th* iterations in SC-MBO- BLS, SC-MBO-KELM, SC-MBO-ELM, and SC-MBO-MLP models

*Kappa*(*Kpa*) =

(1 — *Expected Acc*)

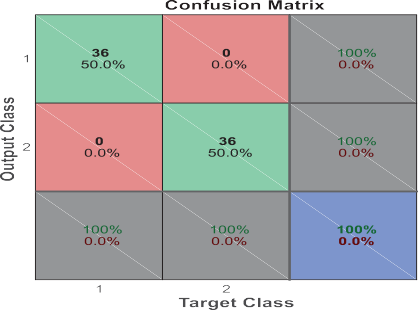
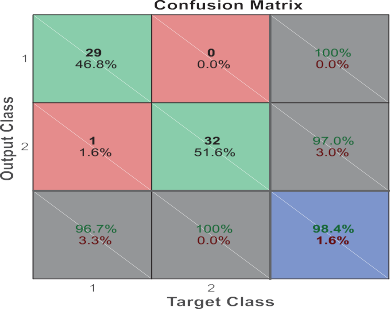
# (37)

respectively. In Ovarian cancer data, the accuracy is converged after 54*th*, 70*th*, 74*th*, and 80*th* iterations in SC-MBO-BLS, SC-MBO-

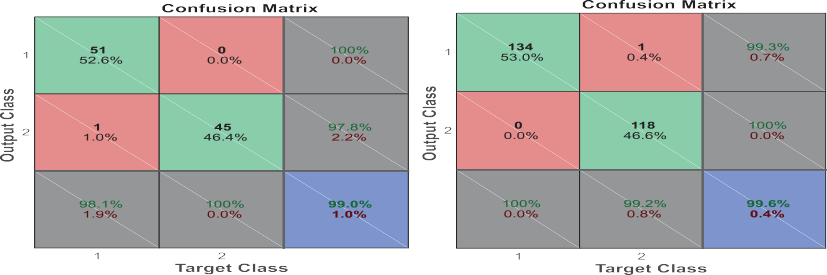
Table 5

Noted performance evaluating parameter’s value (in %) of the SC-MBO-BLS.

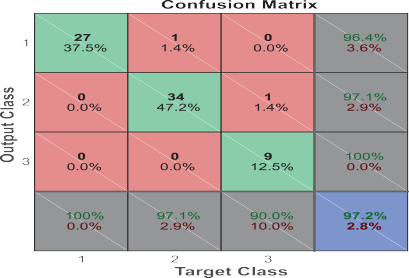
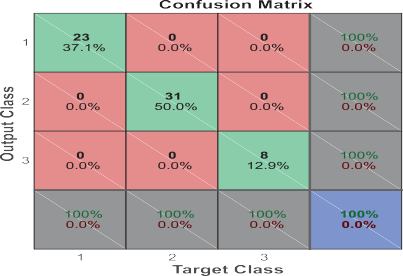
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Types of Datasets | Dataset | Acc | Sen | Spe | MCC | F-measure | Kappa |
| Binary | Leukemia | 100 | 100 | 100 | 100 | 98.83 | 94.32 |
| class | Colon tumor | 98.4 | 96.67 | 100 | 96.82 | 98.31 | 96.77 |
|  | Breast cancer | 99 | 98.08 | 100 | 97.95 | 99.03 | 97.93 |
|  | Ovarian cancer | 99.6 | 100 | 99.16 | 99.21 | 99.63 | 99.21 |
| Multi- | Lymphoma-3 | 100 | 99.97 | 100 | 99.98 | 100 | 100 |
| class | MLL | 97.2 | 96.43 | 97.73 | 94.16 | 96.43 | 94.16 |
|  | ALL-AML-3 | 100 | 99.98 | 100 | 100 | 100 | 100 |
|  | SRBCT | 100 | 100 | 100 | 100 | 100 | 100 |
|  | ALL-AML-4 | 98.6 | 100 | 97.22 | 97.26 | 98.63 | 97.22 |
|  | Lung cancer | 99.5 | 99.09 | 100 | 99.02 | 99.54 | 99.01 |

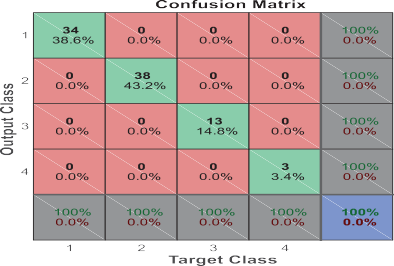
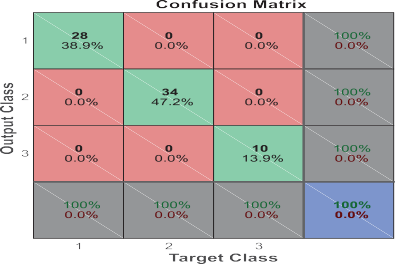
(a) Leukemia (b) Colon



(c) Breast cancer (d) Ovarian



(e) Lymphoma (f) MLL



(g) ALL-MLL-3 (h) SRBCT

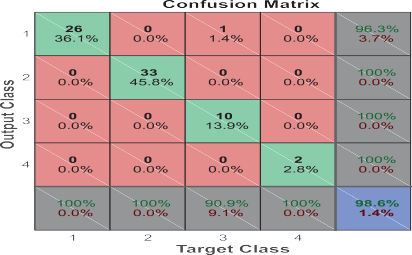
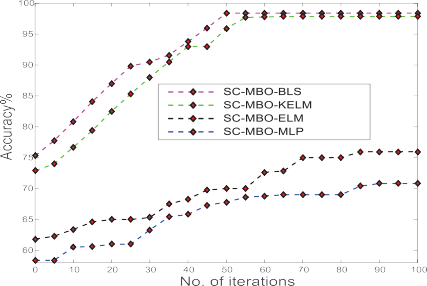
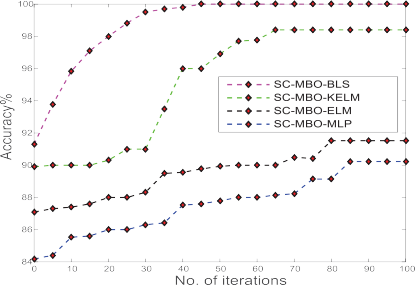
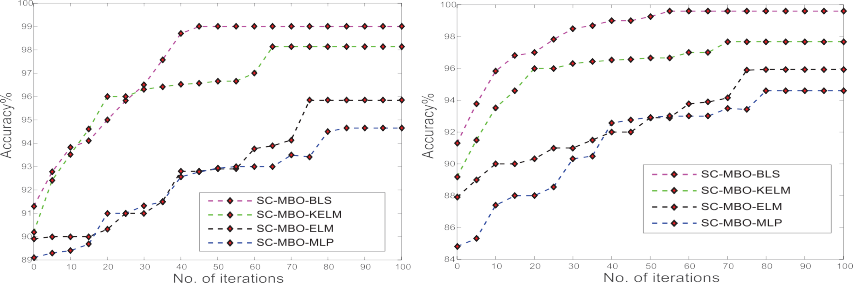


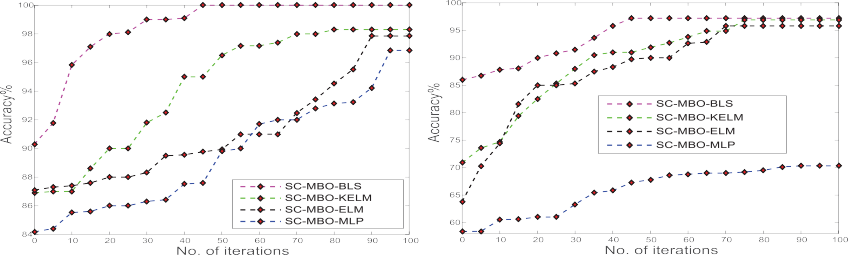
Fig. 6. Confusion matrix of ten microarray dataset.

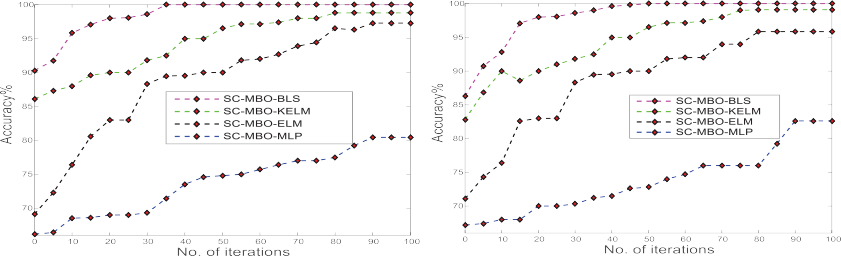
(a) Leukemia (b) Colon



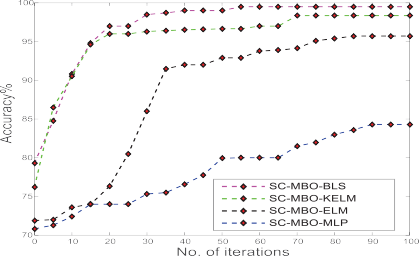
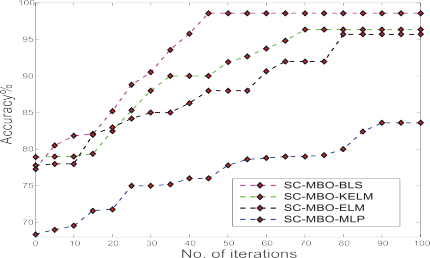
(c) Breast cancer (d) Ovarian



(e) Lymphoma (f) MLL



(g) ALL-MLL-3 (h) SRBCT



(i) ALL-MLL-4 (j) Lung cancer

Fig. 7. Convergence graphs of ten microarray datasets.

Table 6

ACC% comparison between all the approaches.

Table 8

Notable genetic features selected by the SC-MBO-BLS technique in six multi-class

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Datasets | SC-MBO- MLP | SC-MBO- ELM | SC-MBO- KELM | SC-MBO- BLS |  | microarray datasets.    Multi-Class Dataset Name of the Gene # Genes selected |
| Leukemia | 90.23 | 91.52 | 98.4 | 100 |  | Lymphoma-3 GENE1622X 3 |
| Colon tumor | 70.82 | 75.96 | 97.85 | 98.4 |  | GENE2403X |
| Breast cancer Ovarian  cancer | 94.65  95.94 | 95.89  97.68 | 98.12  98 | 99  99.6 |  | GENE2152X  MLL 32847\_ at 4  1389\_ at |

Table 7

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lymphoma-3 | 96.98 | 97.85 | 98.3 | 100 |
| MLL | 70.32 | 95.8 | 96.9 | 97.2 |
| ALL-AML-3 | 80.45 | 97.3 | 98.8 | 100 |
| SRBCT | 82.62 | 95.85 | 99.1 | 100 |
| ALL-AML-4 | 83.6 | 95.7 | 96.34 | 98.6 |
| Lung | 84.3 | 95.72 | 98.35 | 99.5 |

Notable genetic features selected bySC-MBO-BLS technique in four binary class microarray datasets.

|  |  |  |
| --- | --- | --- |
| Dataset | Name of the Gene | # Genes selected |
| Leukemia | Y00787\_s\_at X95735\_at M23197\_at | 3 |
| Colon tumor | M63391 H08393 | 4 |
|  | Z50753  M26383 |  |
| Breast cancer | Contig48393\_RC Contig25534\_RC N65982 AI830996 | 4 |
| Ovarian cancer | MZ244.95245 MZ245.24466 MZ246.41524 | 3 |

KELM, SC-MBO-ELM, and SC-MBO-MLP models respectively. In Lymphoma-3 cancer data, the accuracy is being converged after

45*th*, 75*th*, 86*th*, and 93*rd* iterations in SC-MBO-BLS, SC-MBO- KELM, SC-MBO-ELM, and SC-MBO-MLP models respectively. In

MLL cancer data, the accuracy is converged after 45*th*, 70*th*, 73*th*,

and 84*th* iterations in SC-MBO-BLS, SC-MBO-KELM, SC-MBO-ELM, and SC-MBO-MLP models respectively. In ALL-AML-3 data, the

accuracy is converged after 35*th*, 74*th*, 84*th*, and 90*th* iterations in SC-MBO-BLS, SC-MBO-KELM, SC-MBO-ELM, and SC-MBO-MLP

models respectively. In SRBCT cancer data, the accuracy is being

converged after 44*th*, 76*th*, 80*th*, and 90*th* iterations in SC-MBO- BLS, SC-MBO-KELM, SC-MBO-ELM, and SC-MBO-MLP models

respectively. In ALL-AML-4 cancer data, the accuracy is being con-

verged after 45*th*, 70*th*, 79*th*, and 90*th* iterations in SC-MBO-BLS, SC- MBO-KELM, SC-MBO-ELM, and SC-MBO-MLP models respectively.

In Lung cancer, the accuracy is converging after 53*th*, 64*th*, 83*th*,

and 88*th* iterations in SC-MBO-BLS, SC-MBO-KELM, SC-MBO-ELM, and SC-MBO-MLP models respectively.

Here, the suggested (SC-MBO-BLS) algorithm is compared with the other approaches like SC-MBO hybridized KELM, SC-MBO hybridized ELM, and SC-MBO hybridized MLP, to show a neutral comparison. This comparison is shown in [Table 6](#_bookmark24). In the SC- MBO-BLS model, the discussed microarray dataset like Leukemia, Colon tumor, Breast cancer, Ovarian cancer, Lymphoma-3, MLL, ALL-AML-3, SRBCT, ALL-AML-4 and Lung cancer have performed 90.23%, 70.82%, 94.65%, 95.94%, 96.98%, 70.32%, 80.45%, 82.62%,

83.6%, and 84.3% accuracy respectively.

* 1. *Selected eminent genomic features by SC-MBO-BLS*

|  |  |  |
| --- | --- | --- |
|  | 37539\_ at |  |
| 39931\_ at |
| ALL-AML-3 | X76223\_s\_at | 3 |
|  | X59871\_at |  |
|  | M21624\_at |  |
| SRBCT | gene 714 | 4 |
|  | gene 1003 |  |
|  | gene 2 |  |
|  | gene 554 |  |
| ALL-AML-4 | M23197\_at | 4 |
|  | M92287\_at |  |
|  | M31303\_rna1\_at |  |
| Lung cancer | 37210\_at  36149\_at | 4 |
|  | 37478\_at |  |
|  | 36924\_r\_at |  |
|  | 37210\_at |  |

In [Tables 7](#_bookmark24) and [8](#_bookmark25), the highly influenced genes are noted which are selected by the SC-MBO-BLS algorithm. From Leukemia, 3 nota- ble genes are selected. From colon, 4 notable genes are selected. From Ovarian, 3 notable genes are selected. Likewise, 4 key genes (i.e., Contig48393\_RC, Contig25534\_RC, N65982, AI830996) are chosen from Breast cancer, 3 key genetic features (i.e., GENE1622X, GENE2403X, GENE2152X) are chosen from Lymphoma-3, 4 key genetic features (i.e., 32847\_ at, 1389\_ at, 37539\_ at, 39931\_ at) are chosen from MLL, 3 key genetic features (i.e., M21624\_at, X76223\_s\_at, X59871\_at) are chosen from ALL-AML-3, 4 key genetic features (i.e., gene 2, gene 554, gene 714, gene 1003) are chosen from SRBCT, 4 key genetic features (i.e., M23197\_at, M92287\_at, M31303\_rna1\_at, 37210\_at) are chosen from ALL- MLL-4, and 4 key genetic features (i.e., 36149\_at, 37478\_at, 36924\_r\_at, 37210\_at) are chosen from Lung cancer.

* 1. *Execution period of the SC-MBO-BLS*

In [Table 9](#_bookmark26), the execution time of both the K-FS part and SC- MBO-BLS part are noted down. The execution time of the presented model relies on the population size, size of iterations, size of sam- ples, and fitness function cost. In [Table 9](#_bookmark26), the execution time of both parts in Leukemia, Colon tumor, Breast cancer, Ovarian can- cer, Lymphoma-3, MLL, SRBCT, ALL-AML-3, ALL-AML-4, and Lung

Table 9

Noted Execution Time (ET) of both K-FS part and SC-MBO-BLS part.

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | ET of K-FS | ET of SC-MBO-BLS | Complete ET |
| Leukemia | 0.192 | 168.125 | 168.317 |
| Colon cancer | 0.712 | 156.62 | 157.332 |
| Breast Cancer | 0.718 | 142.62 | 143.338 |
| Ovarian | 0.75 | 58.625 | 59.375 |
| Lymphoma-3 | 0.625 | 60.312 | 60.937 |
| MLL | 0.625 | 143.62 | 144.245 |
| SRBCT | 0.165 | 172.283 | 172.448 |
| ALL-AML-3 | 0.562 | 142.845 | 143.407 |
| ALL-AML-4 | 0.714 | 190.82 | 191.534 |
| Lung cancer | 3.025 | 242.74 | 245.765 |

Table 10

A qualitative ndard models.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Existing approaches | Leukemia | Colon | Breast | Ovarian | Lymphoma-3 | MLL | ALL-AML-3 | SRBCT | ALL-AML-4 | Lung |
| PSO-AKNN [[7]](#_bookmark45) | – | – | – | – | – |  | 90.66 | 94(8.5) | – | – |
|  |  |  |  |  |  |  | (3.3) |  |  |  |
| IWSS-MB-NB [[55]](#_bookmark57) | 97.1 (6.4) | 86 (5.2) | – | – | – |  | – | – | – | – |
| DRFO-CFS [[56]](#_bookmark57) | 91.18 (13) | 90 (10) |  | 100 (16) | – |  | – | – | – | – |
| mRMR-ABC [[9]](#_bookmark29) | – | – |  | – | 96.96 (5) |  | 96.12 (20) | 96.30 (10) | – | – |
| CC-PSO[[58]](#_bookmark57) | – | – |  | – | 96.8(306) |  | 93.7(63) | – | – | – |
| 8-SPMSO [[59]](#_bookmark57) | 98.1(20) | 94.2 (20) |  | – | – |  | – | – | – | – |
| GBC [[8]](#_bookmark29) | – | – | – | – | 98.48(5) |  | 95.83(8) | 96.38(6) |  |  |
| MCSO [[13]](#_bookmark29) | – | – |  | – | – |  | – | 71.04(100) | – | – |
| FRQR [[60]](#_bookmark57) | 97.22 |  | – | 99.60 | – | – | – | – | – | 99.45 |
|  | (7) |  |  | (9) |  |  |  |  |  | (6) |
| GEM [[5]](#_bookmark43) | 91.5 (3) | 91.2 (8) |  | – | – |  | – | – |  |  |
| BDE-XRankf | 82.4(6) | 75(4) |  | 95(3) | – |  | – | – | – | – |
| [[61]](#_bookmark57)  WCSSA-KELM [[62]](#_bookmark57) | 99.0 | 95.5 | 94.3 | 100 | 99.71 | – | 99.38 | 100 | – | 98.90 (5) |
|  | (3) | (5) |  | (4) | (3) |  | (4) | (7) |  |  |
| CBFS-ISVM [[63]](#_bookmark57) | 98.75 | – | – | – | – | 90.53 | – | 98.89 | – | – |
|  | (3) |  |  |  |  | (3) |  | (5) |  |  |
| NB-ERGS [[64]](#_bookmark57) |  | 82.86 |  | 99.49 |  | 88.89 | 93.06 | – | – | 98.34 |
|  |  | (10) |  | (2.8) |  | (10) | (10) |  |  | (10) |
|  |  | - |  |  |  |  |  |  |  |  |
| KNN-IGIS [[65]](#_bookmark57) | – | 77.47 | – | 99.49 | 94.37 | 84.12 | 89.05 | 91.35 | 81.9 | 92.06 |
|  |  | (5.3) |  | (2.8) | (3) | (7) | (4.7) | (8.3) | (6.1) | (12.4) |
| RMA-SVM [[66]](#_bookmark58) | 91.67 | – | – | – | – | 95.83 | – | 97.59 | – | – |
|  | (4) |  |  |  |  | (6) |  | (5) |  |  |
| DR-FS-MFMR [[67]](#_bookmark59) | 31.94[[10]](#_bookmark29) | 88.06 [[40]](#_bookmark52) | – | – | 61.71[90] | – | – | – | – |  |
| FS-DANI [[68]](#_bookmark60) | 93.74 ± 3.21 | – | 62.49 ± 4.31 | 91.68 ± 3.55 | – | – | – | – | – |  |
| RFODL-MGEC [[69]](#_bookmark61) | – | 94.74 | – | 98.68 | – | – | – | – | – |  |
| CDNC-ELM [[70]](#_bookmark63) | 90.18 | 88.73 | – | – | – | – | – | 82.82 | – |  |
| FLAE-OFSP [[71]](#_bookmark64) | 96.09 | 81.83 | – | 99.6 | 89.35 | 92.75 | – | – | – |  |
| SC-MBO-MLP | 90.23 (14) | 70.82 (15) | 94.65 (17) | 95.94 (12) | 96.98 (14) | 70.32 (18) | 80.45 (13) | 82.62 (11) | 83.6 (18) | 84.3 (15) |
| SC-MBO-ELM | 91.52 (11) | 75.96 (14) | 95.89 (13) | 97.68 (16) | 97.85 (11) | 95.8 | 97.3 (16) | 95.85 (9) | 95.7 (12) | 95.72 (17) |
|  |  |  |  |  |  | (15) |  |  |  |  |
| SC-MBO-KELM | 98.4 (12) | 97.85 (10) | 98.12 (10) | 98 (9) | 98.3 | 96.9 (7) | 98.8 (13) | 99.1 (10) | 96.34 (10) | 98.35 (16) |
|  |  |  |  |  | (8) |  |  |  |  |  |
| SC-MBO-BLS | 100 (3) | 98.4 (4) | 99 (4) | 99.6 (3) | 100 (3) | 97.2 (4) | 100 (3) | 100 (4) | 98.6 (4) | 99.5 (4) |

Table 11

ANOVA Test comparison in terms of Acc%.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Dataset | SC-MBO-MLP | SC-MBO-ELM | SC-MBO-KELM | SC-MBO-BLS | Total |
| Number of datasets | 10 | 10 | 10 | 10 | 40 |
|  | 849.91 | 939.27 | 980.16 | 992.3 | 2314.84 |
| Mean | 84.991 | 93.927 | 98.016 | 99.23 | 57.871 |
|  | 89627.8639 | 92510.6243 | 96203.9338 | 98473.77 | 223434.3422 |
| Standard deviation | 0.303242317 | 0.8578973 | 0.886303184 | 0.933392855 | 0.745208914 |

PX PX2

Table 12

P-value estimation using anova test.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | SS | Df | MS |  |
| Between-treatments | 1246.41842 | 3 | 415.4728067 | *f* value = 12.09611 |
| Within-treatments | 1236.515184 | 36 | 34.347644 | *p* value = 0.030957 |
| Total | 2482.934 | 39 |  |  |

cancer are 168.317, 157.332, 143.338, 59.375, 60.937, 144.245,

172.448, 143.407, 191.534, and 245.765 respectively.

* 1. *Qualitative analysis*

A qualitative based analysis has been made to show a neutral comparison. In [Table 10](#_bookmark27), 16 standard models have been compared with K-FS based SC-MBO-BLS technique by its classification accu- racy and selected genetic features.

From the above table, it is revealed that K-FS based SC-MBO-BLS performs better than others. Some model like DRFO-CFS and WCSSA-KELM [[62]](#_bookmark57) achieve 100% accuracy in Ovarian and SRBCT respectively but these models select a greater number of gene as compared to the proposed method.

* 1. *A Statistic-based comparison in terms of classification accuracy*

To find the mean value of the definite groups, a widely approved statistical evaluator i.e., ANOVA (Analysis of variance)

has been taken in this research experiment. ANOVA performs a sta- tistical evaluation of the presented method. Usually, ANOVA is based on a null hypothesis. In this evaluating approach, initially, the F-value is computed, then the p-value is observed as per F- value. According to the p-value, it is finalized whether to observe or discard the alternative hypothesis. If the p-value is less than or equal to 0.05 (considering the significance level to 5%) then the null hypothesis is rejected, and the accuracies of all the tech- niques are dissimilar. Eventually, the statistical evaluation of the model through the ANOVA test is given in [Tables 11 and 12](#_bookmark28). Here, the p-value is found as 0.0309. this value is smaller than the for- merly taken p-value. So, the alternative hypothesis is dismissed and the presented method is more statistically significant than others.

1. Conclusion

In this presented research study, an optimization algorithm Sine-Cosine hybridized MBO is merged with an eminent classifier, viz., BLS to choose the most notable genetic features with enhanced classification results. In this study, to estimate the pre- sented algorithm, ten microarray datasets are considered in which 4 datasets are binary and the other six are multiclass. In the first stage, a pre-extraction technique (Kernel-based FS) is applied to choose subsets of genes and then these extracted genetic feature subsets have to go through the SC-MBO-BLS model for further eval- uation. Here, other performance evaluators like sensitivity, preci- sion, specificity, F-score, Kappa, and MCC are taken for a neutral comparison. Moreover, to show the dominance of the suggested technique, some standard approaches like SC-MBO-MLP, SC\_MBO-ELM, SC\_MBO-KELM, and 16 standard models have been compared with K-FS based SC-MBO-BLS technique by its classifica- tion accuracy and selected genetic features. Eventually, to compute the mean value of the definite groups, a widely approved statistical evaluator i.e., ANOVA (Analysis of variance) has been taken in this research experiment. From the above statistical and quantitative study, it is concluded that the presented SC-MBO-BLS technique will be a trustworthy support for the detection of numerous diseases.

Declaration of Competing Interest

The authors declare that they have no known competing finan- cial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. [Aydadenta H, Adiwijaya A. A clustering approach for feature selection in](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0005) [microarray data classification using random forest. J Inf Process Syst 2018;14](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0005) [(5):1167–75](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0005).
2. [Wang H, Jing X, Niu B. A discrete bacterial algorithm for feature selection in](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0010) [classification of microarray gene expression cancer data. Knowl-Based Syst](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0010) [2017;126:8–19](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0010).
3. [Bicciato S, Luchini A, Di Bello C. PCA disjoint models for multiclass](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0015) [cancer analysis using gene expression data. Bioinformatics 2003;19](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0015) [(5):571–8](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0015).
4. [Ang JC, Mirzal A, Haron H, Hamed HNA. Supervised, unsupervised, and semi-](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0020) [supervised feature selection: a review on gene selection. IEEE/ACM Trans](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0020) [Comput Biol Bioinform (TCBB) 2016;13(5):971–89](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0020).
5. [Hernandez JCH, Duval B, Hao JK. A genetic embedded approach for gene](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0025) [selection and classification of microarray data. In: European Conference on](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0025) [Evolutionary Computation, Machine Learning and Data Mining in](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0025) [Bioinformatics. Springer; 2007. p. 90–101](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0025).
6. [Shukla AK, Singh P, Vardhan M. A two-stage gene selection method for](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0030) [biomarker discovery from microarray data for cancer classification. Chemom](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0030) [Intel Lab Syst 2018;183:47–58](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0030).
7. [Kar S, Sharma KD, Maitra M. Gene selection from microarray gene](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0035) [expression data for classification of cancer subgroups employing PSO and](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0035) [adaptive K-nearest neighborhood technique. Expert Syst Appl 2015;42](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0035) [(1):612–27](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0035).
8. [Alshamlan HM, Badr GH, Alohali YA. Genetic bee colony (GBC) algorithm: a](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0040) [new gene selection method for microarray cancer classification. Comput Biol](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0040) [Chem 2015;56:49–60](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0040).
9. H. Alshamlan, G. Badr, Y. Alohali, mRMR-ABC: A hybrid gene selection algorithm for cancer classification using microarray gene expression profiling, BioMedRes.Int.2015(2015)604910–604910.
10. Choubey, Dilip Kumar, et al. ‘‘Classification of Pima indian diabetes dataset using naive bayes with genetic algorithm as an attribute selection.” *Communication and computing systems: proceedings of the international conference on communication and computing system (ICCCS 2016)*. 2017.
11. [Akizur RM, Muniyandi RC. Feature selection from colon cancer dataset for](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0055) [cancer classification using artificial neural network. Int J Adv Sci Eng Inform](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0055) [Technol 2018;8(4-2):1387–93](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0055).
12. [Rodrigues D et al. A wrapper approach for feature selection based on bat](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0060) [algorithm and optimum-path forest. Expert Syst Appl 2014;41(5):2250–8](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0060).
13. [Mohapatra P, Chakravarty S, Dash PK. Microarray medical data classification](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0065) [using kernel ridge regression and modified cat swarm optimization-based](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0065) [gene selection system. Swarm Evol Comput 2016;28:144–60](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0065).
14. [Rani MJ, Devaraj D. Two-stage hybrid gene selection using mutual information](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0070) [and genetic algorithm for cancer data classification. J Med Syst 2019;43](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0070) [(8):1–11](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0070).
15. [Cristianini N, Shawe-Taylor J. An introduction to support vector machines and](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0075) [other kernel-based learning methods. Cambridge University Press; 2000](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0075).
16. [Ibrahim HT et al. A grasshopper optimizer approach for feature selection and](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0080) [optimizing SVM parameters utilizing real biomedical data sets. Neural Comput](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0080) [& Applic 2019;31(10):5965–74](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0080).
17. [Malathi V, Marimuthu NS, Baskar S. Intelligent approaches using support](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0085) [vector machine and extreme learning machine for transmission line](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0085) [protection. Neurocomputing 2010;73(10-12):2160–7](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0085).
18. [Zurada JM. Introduction to artificial neural systems, Vol. 8. St. Paul: West;](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0090) [1992](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0090).
19. [Aydogan EK, Karaoglan I, Pardalos PM. HGA: hybrid genetic algorithm in fuzzy](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0095) [rule-based classification s high-dimensional problems. Appl Soft Comput](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0095) [2012;12(2):800–6](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0095).
20. [Anagaw A, Chang Y-L. A new complement naïve Bayesian approach for](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0100) [biomedical data classification. J Ambient Intell Hum Comput 2019;10](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0100) [(10):3889–97](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0100).
21. [Naik B et al. A harmony search based gradient descent learning-FLANN (HS-](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0105) [GDL-FLANN) for classification. In: Computational Intelligence in Data Mining-](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0105) [Volume 2. New Delhi: Springer; 2015. p. 525–39](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0105).
22. [Heermann PD, Khazenie N. Classification of multispectral remote sensing data](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0110) [using a back-propagation neural network. IEEE Trans Geosci Remote Sens](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0110) [1992;30(1):81–8](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0110).
23. Al-Shargabi, Bassam, Feda Alshami, and Rami Alkhawaldeh. ‘‘Enhancing multi- layer perception for breast cancer prediction.”*International Journal of Advanced Science and Technology*(2019).
24. [Fernández-Navarro F et al. Evolutionary generalized radial basis function](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0120) [neural networks for improving prediction accuracy in gene classification using](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0120) [feature selection. Appl Soft Comput 2012;12(6):1787–800](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0120).
25. [Guliyev NJ, Ismailov VE. A single hidden layer feedforward network with only](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0125) [one neuron in the hidden layer can approximate any univariate function.](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0125) [Neural Comput 2016;28(7):1289–304](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0125).
26. [Arulampalam G, Bouzerdoum A. A generalized feedforward neural network](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0130) [architecture for classification and regression. Neural Netw 2003;16(5-](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0130) [6):561–8](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0130).
27. [Pao Y-H, Takefuji Y. Functional-link net computing: theory, system](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0135) [architecture, and functionalities. Computer 1992;25(5):76–9](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0135).
28. Samanta, Sourav, et al. ‘‘Haralick features based automated glaucoma classification using back propagation neural network.”*Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014*. Springer, Cham, 2015.
29. [Igelnik B, Pao Y-H. Stochastic choice of basis functions in adaptive function](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0145) [approximation and the functional-link net. IEEE Trans Neural Netw 1995;6](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0145) [(6):1320–9](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0145).
30. [Huang G-B, Chen Y-Q, Babri HA. Classification ability of single hidden](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0150) [layer feedforward neural networks. IEEE Trans Neural Netw 2000;11](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0150) [(3):799–801](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0150).
31. LeCun, Yann, et al. ‘‘Handwritten digit recognition with a back-propagation network.”*Advances in neural information processing systems*2 (1989).
32. Denker, John S., et al. ‘‘Neural network recognizer for hand-written zip code digits.”*Advances in neural information processing systems*. 1989.
33. [Shen Lu et al. Multiple empirical kernel mapping based broad learning system](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0165) [for classification of Parkinson’s disease with transcranial sonography. 2018](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0165) [40th Annual International Conference of the IEEE Engineering in Medicine and](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0165) [Biology Society (EMBC). IEEE; 2018](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0165).
34. [Pao Y-H, Park G-H, Sobajic DJ. Learning and generalization characteristics](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0170) [of the random vector functional-link net. Neurocomputing 1994;6](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0170) [(2):163–80](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0170).
35. [Chen CLP, Zhang C-Y. Data-intensive applications, challenges, techniques and](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0175) [technologies: a survey on Big Data. Inf Sci 2014;275:314–47](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0175).
36. [Chen CLP, Liu Z. Broad learning system: An effective and efficient incremental](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0180) [learning system without the need for deep architecture. IEEE Trans Neural](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0180) [Networks Learn Syst 2017;29(1):10–24](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0180).
37. [Lu H et al. A kernel extreme learning machine algorithm based on improved](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0185) [particle swam optimization. Memetic Comput 2017;9(2):121–8](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0185).
38. [Sayyad H, Manshad AK, Rostami H. Application of hybrid neural particle](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0190) [swarm optimization algorithm for prediction of MMP. Fuel](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0190) [2014;116:625–33](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0190).
39. [Mansour IB, Alaya I, Tagina M. A gradual weight-based ant colony approach for](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0195) [solving the multiobjective multidimensional knapsack problem. Evol Intel](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0195) [2019;12(2):253–72](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0195).
40. [Zhang H et al. Developing a novel artificial intelligence model to estimate the](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0200) [capital cost of mining projects using deep neural network-based ant colony](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0200) [optimization algorithm. Resour Policy 2020;66:101604](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0200).
41. [Mohapatra P, Chakravarty S, Dash PK. An improved cuckoo search based](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0205) [extreme learning machine for medical data classification. Swarm Evol Comput](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0205) [2015;24:25–49](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0205).
42. [Yamany W et al. Moth-flame optimization for training multi-layer](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0210) [perceptrons. 2015 11th International computer engineering Conference](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0210) [(ICENCO). IEEE; 2015](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0210).
43. [Majhi SK. How effective is the moth-flame optimization in diabetes data](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0215) [classification. In: Recent Developments in Machine Learning and Data](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0215) [Analytics. Singapore: Springer; 2019. p. 79–87](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0215).
44. [Chang Y-T et al. Optimization the initial weights of artificial neural networks](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0220) [via genetic algorithm applied to hip bone fracture prediction. Adv Fuzzy Syst](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0220) [2012;2012](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0220).
45. [Li H et al. Genetic algorithm for the optimization of features and neural](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0225) [networks in ECG signals classification. Sci Rep 2017;7(1):1–12](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0225).
46. [Wang G-G, Deb S, Cui Z. Monarch butterfly optimization. Neural Comput &](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0230) [Applic 2019;31(7):1995–2014](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0230).
47. [Mirjalili S. SCA: a sine cosine algorithm for solving optimization problems.](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0235) [Knowl- Based Syst 2016;96:120–33](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0235).
48. [Polat K, Güne](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0240) s [S. A new feature selection method on classification of medical](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0240) [datasets: Kernel F-score feature selection. Expert Syst Appl 2009;36](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0240) [(7):10367–73](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0240).
49. [Kira K, Rendell LA. A practical approach to feature selection. In: Machine](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0245) [learning proceedings 1992. Morgan Kaufmann; 1992. p. 249–56](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0245).
50. [Cai H et al. Feature weight estimation for gene selection: a local hyperlinear](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0250) [learning approach. BMC Bioinf 2014;15(1):1–13](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0250).
51. [Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al.](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0255) [Molecular classification of cancer: class discovery and class prediction by gene](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0255) [expression monitoring. Science 1999;286(5439):531–7](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0255).
52. [Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, et al. Broad patterns](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0260) [of gene expression revealed by clustering analysis of tumor and normal colon](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0260) [tissues probed by oligonucleotide arrays. Proc Natl Acad Sci 1999;96](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0260) [(12):6745–50](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0260).
53. [http://csse.szu.edu.cn/staff/zhuzx/Datasets.html.](http://csse.szu.edu.cn/staff/zhuzx/Datasets.html)
54. [Petricoin III EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al.](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0270) [Use of proteomic patterns in serum to identify ovarian cancer. Lancet](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0270) [2002;359(9306):572–7](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0270).
55. [Zhu Z, Ong Y-S, Dash M. Markov blanket-embedded genetic algorithm for gene](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0275) [selection. PatternRecognit 2007;40(11):3236–48](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0275).
56. [Wang A et al. Incremental wrapper based gene selection with Markov blanket.](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0280) [2014 IEEE International Conference on Bioinformatics and Biomedicine](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0280) [(BIBM). IEEE; 2014](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0280).
57. [Bolón-Canedo V, Sánchez-Maroño N, Alonso-Betanzos A. Distributed feature](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0285) [selection: An application to microarray data classification. Appl SoftComput](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0285) [2015;30:136–50](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0285).
58. Chinnaswamy, Arunkumar, and Ramakrishnan Srinivasan. ‘‘Hybrid feature selection using correlation coefficient and particle swarm optimization on microarray gene expression data.”*Innovations in bio-inspired computing and applications*. Springer, Cham, 2016. 229-239.
59. [García-Nieto J, Alba E. Parallel multi-swarm optimizer for gene selection in](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0295) [DNA microarrays. Appl Intell 2012;37(2):255–66](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0295).
60. [Arunkumar C, Ramakrishnan S. Attribute selection using fuzzy roughset based](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0300) [customized similarity measure for lung cancer microarray gene expression](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0300) [data. Future Comput Inf J 2018;3(1):131–42](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0300).
61. [Apolloni J, Leguizamón G, Alba E. Two hybrid wrapper-filter feature selection](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0305) [algorithms applied to high-dimensional microarray experiments. Appl Soft](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0305) [Comput 2016;38:922–32](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0305).
62. [Baliarsingh SK et al. Analysis of high-dimensional genomic data employing a](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0310) [novel bio-inspired algorithm. Appl Soft Comput 2019;77:520–32](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0310).
63. [Maulik U, Chakraborty D. Fuzzy preference based feature selection and](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0315) [semisupervised SVM for cancer classification. IEEE Trans NanoBiosci 2014;13](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0315) [(2):152–60](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0315).
64. [Chandra B, Gupta M. An efficient statistical feature selection approach for](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0320) [classification of gene expression data. J Biomed Inform 2011;44(4):529–35](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0320).
65. [Nakariyakul S. A hybrid gene selection algorithm based on interaction](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0325) [information for microarray-based cancer classification. PLoS One 2019;14(2):](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0325) [e0212333](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0325).
66. [Ghosh M et al. Recursive memetic algorithm for gene selection in microarray](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0330) [data. Expert Syst Appl 2019;116:172–85](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0330).
67. [Saberi-Movahed F et al. Dual regularized unsupervised feature selection based](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0335) [on matrix factorization and minimum redundancy with application in gene](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0335) [selection. Knowl-Based Syst 2022;256:109884](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0335).
68. [Qi Y et al. A new feature selection method based on feature distinguishing](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0340) [ability and network influence. J Biomed Inform 2022;128:104048](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0340).
69. [Vaiyapuri T et al. Red fox optimizer with data-science-enabled microarray](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0345) [gene expression classification model. Appl Sci 2022;12(9):4172](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0345).
70. [Rostami M et al. Gene selection for microarray data classification via multi-](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0350) [objective graph theoretic-based method. Artif Intell Med 2022;123:102228](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0350).
71. [Nosrati V, Rahmani M. An ensemble framework for microarray data classification](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0355) [based on feature subspace partitioning. Comput Biol Med 2022;148:105820](http://refhub.elsevier.com/S1110-8665(23)00019-1/h0355).