

Modified Nelder-Mead Method in Microarray Data using Bi-Clustering

Anil Kumar R J¹, Veena M N², Niirmala M S³ and Nagendra Nath Giri⁴

¹⁻⁴Maharanis Science College for Women PESCE Mandya Government College for Women Maharanis Science for Women Mysore
Email: anilkumar.rj@gmail.com¹, veenadishal@gmail.com², nimsms@gmail.com³ and nagengiri123@gmail.com⁴

Received:03/08/2025

Revised: 18/08/2025

Accepted:08/09/2025

Published:25/09/2025

ABSTRACT

The Modified Nelder-Mead algorithm for biclustering microarray gene expression data is proposed to overcome the poor convergence problem of NM method. It focuses on finding coherent biclusters with lower MSR and higher row variance. In Nelder-Mead method the median is measured instead of mean. The median provides much better estimates in place of mean. Before shrinking operation, the differential evolution is applied to obtain global minimal solution. A qualitative measure of the formed biclusters with a comparative assessment of results are provided on two benchmark gene expression datasets to demonstrate the effectiveness of the proposed method. Biological validation of the selected genes within the biclusters is provided by publicly available GO consortium. The patterns present a significant biological relevance in terms of related biological processes, components and molecular functions in a species-independent manner. In conclusion, it is found that the Modified Nelder-Mead approach gives a better result over the conventional Nelder-Mead method and existing biclustering algorithms.

Keywords: Microarray Gene Expression Data, Cancer Classification, Feature Selection, Differential Evolution (DE), Hybrid Filter-DE Algorithm.



© 2025 by the authors; licensee Advances in Consumer Research. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BYNC.ND) license(<http://creativecommons.org/licenses/by/4.0/>).

INTRODUCTION

Cancer research is widely acknowledged as a highly promising domain for using machine learning. Extensive endeavours have been undertaken to explore prospective approaches for detecting and treating cancer [1].

Cancer is a condition that exhibits uncontrolled cellular proliferation and results as a growth of a tumor in the form of mass or lump. Lung, colon, breast, central nervous system (CNS), liver, kidney, prostate, and brain cancer are among the various types of cancer that can occur. In this research study we have examined four distinct types of cancer dataset: Lung, Breast, Brain, and Central Nervous System. Lung cancer is a prevalent and mortal cancer worldwide [2]. It can arise in the primary airway, specifically within the lung tissue. The outcome is the unregulated proliferation and growth of specific lung cells. Respiratory disorders, including emphysema, are linked to an increased risk of lung cancer development. Breast cancer is one of the most invasive malignancies, predominantly affecting women. It is considered the most severe cancer following lung cancer due to the elevated mortality rate among women [3,4]. The rapid development of abnormal brain cells that is indicative of a brain tumor [5,6,7] is a significant health concern for adults, as it can result in severe impairment of organ function and even mortality. A malignant brain tumour rapidly grows and extends to

adjacent brain regions. The Central Nervous System (CNS), consisting of the brain and spinal cord, is responsible for numerous biological functions. Spinal cord compression and spinal instability often involve the vertebral and spinal epidural spaces as common sites for cancer metastases. Metastases represent the most common type of CNS tumour in adults [8].

Cancer is regarded as one of the primary causes of death. In order to preserve the lives of patients, advanced technologies such as artificial intelligence and machine learning are used to detect cancer at an early stage and accurately predict its type. The cancer diagnosis is performed by employing several medical datasets, which encompass microarray gene expression data, also known as the microarray dataset. Microarray technology offers unique experimental capabilities that have been beneficial to cancer research. Microarray data can be used to evaluate a wide variety of cancer types. High-dimensional data from DNA microarray experiments is known as gene expression data. It is widely used to classify and detect malignant disorders [9]. The most recent development of artificial intelligence, specifically machine learning, has simplified data analysis, including microarray data. The authors [10] demonstrated that machine learning algorithms can be employed for microarray dataset analysis for cancer classification. Utilizing expressions of genes in microarray datasets can serve as an effective

tool for diagnosing cancer. However, the number of active genes continues to grow, surpassing hundreds of thousands, while the available datasets remain limited in size, containing only a few subsets of samples. Therefore, one of the challenges in analyzing microarray datasets used for cancer classification is the curse of dimensionality. There is an additional concern regarding the characteristics of the current microarray datasets, which consist of numerous redundant and irrelevant features that have a detrimental impact on cancer classification results and computational expense [11]. The presence of duplicated and irrelevant features in very high-dimensional microarray datasets reduces the ability of the machine learning techniques to achieve accurate cancer classification and prediction [12]. These characteristics diminish the efficiency of the prediction model and complicate the search for meaningful insights. Consequently, it is necessary to employ feature selection methods in order to enhance the accuracy of the machine learning classifiers [13]. In order to enhance the effectiveness of widely used machine learning algorithms, many feature selection techniques have been employed to identify the most important features in malignant microarray datasets [14,15,16,17,18]. Even though filter feature selection approaches offer computational efficiency and the ability to reduce the dimensionality of microarray datasets, their accuracy results are limited since they evaluate features independently of classifiers. On the other hand, wrapper feature selection approaches interact with the classifier throughout the feature evaluation process, resulting in superior outcomes compared to the filter method. Nevertheless, the utilization of wrapper approaches on high-dimensional microarray datasets might be difficult and time-consuming.

In recent years, several evolutionary and bio-inspired algorithms [19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35] have been implemented in literature to obtain the highest level of accuracy in the gene selection challenge. Although feature selection methods based on evolutionary algorithms can overcome the limitations of filter and wrapper methods, they may result in greater computational times for certain machine learning algorithms. Due to the high dimensionality and large number of features in malignant microarray datasets, it is not feasible to initially employ evolutionary algorithms as feature selection approaches. It is essential to reduce the features of microarray cancer datasets using filter feature selection. Then, an evolutionary optimization algorithm can be utilized to optimize the features further to maximize cancer classification performance. This motivated us to suggest a novel hybrid filter-differential evolutionary feature selection method that combines the strengths of both filters and evolutionary techniques to generate effective solutions with improved cancer classification performance for high-dimensional microarray datasets.

The Differential Evolutionary (DE) is one of the superior optimization evolutionary algorithms, which is inspired by the biological evolution of the chromosomes

in nature. DE performs well in convergence, although it is straightforward to implement and requires a few parameters to control and low space complexity. These attractive advantages of DE over other competitive optimization algorithms make DE gained widespread recognition for its exceptional efficacy in addressing various optimization challenges. Hence, this study aims to combine the superior performance of the DE optimization algorithm with filter selection methods to improve the classification accuracy of four microarray datasets by highlighting the most important and relevant genes. This is the first attempt at applying the hybrid filter and DE-based gene selection and classification of DNA Microarray data to the belief of our knowledge. In this paper, we propose a novel approach that combines feature selection methods based on differential evolutionary optimization algorithms and filter methods for identifying the most effective subset of features. Six common filtering methods were applied in this study to assign a score to each feature in microarray cancer datasets. These methods were then used to reduce the dimensionality of the datasets by retaining only the highest-ranked features and removing superfluous and irrelevant ones. The DE algorithm was then used to optimize the reduced cancer datasets, resulting in significantly improved results in cancer classification. Our proposed approach improved the classification performance of cancer when applied to microarray datasets with high dimensions.

The remaining part of this paper consists of the following sections. Section 2 discussed recent works related to cancerous gene selection and classification performance suggested for high-dimensional microarrays. Section 3 describes proposed methodology and elaborates on the details of the phases of the proposed hybrid filter-DE feature selection methods. The experimental results and discussion of the proposed hybrid filter-DE are presented in Section 4. Finally, the paper is concluded, and future work is recommended in Section 5.

RELATED WORKS

This section presents and investigates the existing hybrid feature selection methods approaches that have recently been applied on cancer microarray datasets to improve cancer classification results.

Karthika et al. [20] employed the mixture model (MM) in addition to the Fast Fourier Transform (FFT) on Microarray Gene Expression data for dimensionality reduction. In order to select an effective feature, they employed optimization techniques called Dragonfly. Nonlinear Regression, DT, RF, and SVM were used as classifiers in this study. The classifiers' performance is evaluated both with and without feature selection methods. Finally, hyper-parameter tuning techniques such as Adaptive Moment Estimation (Adam) and Random Adaptive Moment Estimation (RanAdam) are used to improve classifiers, resulting in an accuracy of approximately 98% with the SVM classifier. This research did not address computational complexity and

model validation techniques. In addition, this study has notable limitations, including population-specific findings, reliance on MAGE data, and the influence of outliers.

Elbashir et al. [21]. suggested a graph attention network (GAT) model to utilize diverse mRNA and miRNA for the prediction of the survival rate of non-small cell lung cancer (NSCLC) using multi-omics data. Chi-square analysis was used to select the most significant features to include in the model. They used the synthetic minority oversampling method (SMOTE) to make the dataset and the concordance index (C-index) more equal, and they tested the model on different sets of omics data. When using combined mRNA and miRNA data, they obtained the highest value of the C-index (0.82) along with the accuracy of 0.75. Chi-Square approaches cannot be regarded as the most ideal feature selection method for highly complicated and correlated biological data; this is a significant limitation of the current research.

Zamri et al. [22] presented a hybrid metaheuristics optimization-based two-stage feature selection model. The SKF-MUT simulated Kalman filter was used in this study to pick microarray features that would make the ANN classification more accurate. The experimental results were validated using eight binary and multiclass benchmark datasets. SKF-MUT effectively selected the correct number of features and achieved 95–100% classification accuracy. The significant limitations of this study include model evaluation relying just on accuracy. Instead, other metrics like precision, recall, F1-score, or AUC-ROC might better assess the model's performance along with accuracy. Further, the computational cost of feature selection has not been discussed.

Ali et al. [23] presented a hybrid filter-genetic feature selection method to reduce microarray dataset dimensionality. The first part of this work used three filter methods: information gain (IG), information gain ratio (IGR), and Chi-squared (CS) to pick the most relevant microarray dataset features. The second phase used a genetic algorithm to optimize the features selected in the first phase of the proposed approach. The proposed method was validated utilizing breast, lung, CNS, and brain cancer microarray datasets. Experimental results indicated the suggested model improved performance of various common machine learning approaches in terms of Accuracy, Recall, Precision, and F-measure and the reported accuracy ranges from 92 to 100%. The limitations of the existing work can be included as computational cost of the feature selection process and also statistical validation not discussed.

Elemam and Elshrkawey [24] introduced a two-stage hybrid feature selection. They began by using feature evaluation methods that included chi-squared, F-statistics, and mutual information (MI) filters. In the second phase, they employed wrapper-based sequential

forward selection with ML models like SVM, DT, RF, and KNN classifiers to find the optimal set of features. The model was then rigorously tested and validated using lung cancer, ovarian cancer, leukemia, and SRBCT datasets. The results were impressive, with an accuracy rate of almost 100 percent and a minimal number of selected features. However, the study's performance was solely measured through accuracy, and the issue of feature redundancy was not adequately addressed. No statistical tests were conducted for model validation, which are the limitations of the existing work. In a recent study, Abasabadi et al. [25] proposed a novel hybrid feature selection method to address the challenge of high dimensionality in microarray datasets. The methodology combines a filter approach (SLI- γ) with a genetic algorithm (GA). In the initial phase, 99% of irrelevant features were eliminated using SLI- γ . The second phase involved the GA optimization of the remaining relevant features to enhance classification accuracy. The results of this method were not only enhanced performance but also a significant reduction in execution time, which is a remarkable achievement. However, the inherent computational complexity associated with GA-based optimization remains a challenge, especially as the dimensionality of datasets increases.

Almutiri et al. [26] proposed a hybrid feature selection method, GI-SVM-RFE, to improve classification accuracy in high-dimensional microarray datasets. The methodology combines the Gini index and SVM-RFE to select informative genes recursively. The results showed enhanced classification accuracy reported as 90.67 compared to other methods without feature selection or using only the Gini index or SVM-RFE. The model not validated statistically.

Similarly, Xie et al. [27] proposed the Multi-Fitness RankAggreg Genetic Algorithm (MFRAG). The methodology employed a genetic algorithm framework to integrate nine feature selection techniques. It uses an ensemble model to assess fitness and guide the evolutionary process. The results indicated that MFRAG demonstrated exceptional performance, achieving an accuracy between 87 and 100 percent, with increased classification accuracy using fewer selected characteristics. The limitations of this study include the potential for overfitting despite the use of the ensemble method and the absence of statistical discussion of model validation.

Dash et al. [28] proposed a hybrid methodology for feature reduction utilizing harmony search and Pareto optimization. The authors employed the Harmony Search algorithm and Gene Selection (AHSGS) to identify the top 100 gene characteristics while also utilizing Bi-objective Pareto optimization to eliminate insignificant gene features. The model was assessed using four publicly available microarray datasets. In all instances, SVM surpassed other classifiers, attaining nearly 100 percent accuracy, with the exception of the Colon dataset, where ANN reached 82 percent

accuracy. The existing work exhibits notable limitations, particularly in the statistical analysis, as the results concerning significance levels are absent. Furthermore, the author failed to address the criteria or methodology employed to ascertain the Harmony Memory Consideration Rate and Pitch Adjusting Rate. Almutiri et al. [29] suggested a fusion-based feature selection framework aimed at mitigating high dimensionality and enhancing classification performance in gene expression microarray data. The framework utilizes a three-layer approach. The first layer has independent feature selection methods for gene ranking and scoring. The second layer consists of a threshold-based filtering step and a final decision layer employing majority or consensus voting. Experiments were conducted on five microarray datasets using an SVM classifier. The results revealed enhanced classification accuracy, achieving up to 97% on the Prostate dataset, alongside dimensionality reduction in comparison to existing methods. The primary limitations of this study are threshold sensitivity and dependence on voting strategy.

Kilicarslan et al. [30] proposed a hybrid model to significantly improve cancer diagnosis. The methodology combined relief and stacked autoencoders for dimension reduction. Then, SVM and CNN were used to improve classification accuracy. The proposed method achieved the highest classification accuracies (98.6%, 99.86%, and 83.95%, respectively) on three microarray datasets (Ovarian, Leukemia, and CNS), outperforming SVM and other tested approaches. The study highlighted the effectiveness of dimension reduction in enhancing classification accuracy. However, this study has most notable limitations as limited comparison with other Feature Selection Methods and process of hyperparameter optimization for the CNN model.

Baliarsingh et al. [31] proposed a microarray-based hybrid cancer classification model. The methodology utilized ANOVA to select relevant genes. Then, the enhanced Jaya (EJaya) algorithm and the forest optimization algorithm (FOA) were utilized to find the best gene subset, and SVM was used for classification. The proposed method reduced features and exceeded benchmark methods in classification accuracy from 96 to 100%. The significant limitation of this study is the use of a single classifier (SVM), which may not generalize across datasets. In addition Parameter tuning may also affect EJaya and FOA algorithm performance. In this study, Almutiri et al. [32] evaluated and compared contemporary hybrid approaches combining bio-inspired evolutionary algorithms for gene selection and cancer classification. The methodology, which was conducted with utmost thoroughness, involved reviewing various algorithms, with a focus on genetic algorithms (GA) as wrapper methods for gene selection. The results revealed that GA is the most extensively used and achieved the highest accuracy with a minimal number of selected genes ranging from 93 to 100%. In contrast, the Firefly

algorithm has not been used as a wrapper approach. The limitation of existing work is the inadequate investigation of alternative hybrid algorithms.

Sayed et al. [33] proposed this study to investigate the efficacy of a Nested Genetic Algorithm (Nested-GA) for feature selection in high-dimensional colon cancer microarray datasets. The methodology used a t-test to preprocess data and a nested approach with two Genetic Algorithms. The outer Genetic Algorithm (OGA-SVM) is used for gene expression data, and the Inner Genetic Algorithm (IGA-NNW) is utilized for DNA methylation data. The validation was performed using five cross-folds, ensuring a thorough examination of the results. Nested-GA outperformed KNN and RF on the colon cancer dataset with 99.9% classification accuracy. This study's main limitation is comparing Nested-GA to a limited set of feature selection algorithms (KNN and RF). A more extensive comparison with additional contemporary methods could yield more significant insights into its performance which is one of the limitations of this work.

Similarly, Ghosh et al. [34] introduced a novel two-stage hybrid model that integrates multiple filter methods with a genetic algorithm (GA) for cancer detection in microarray datasets. The methodology involved initially creating an ensemble of filter methods such as ReliefF, chi-square, and symmetrical uncertainty by looking at the union and intersection of their top-n-ranked features. Then, in the next step, GA is used to make the results of the first step even better. The result showed that the model did better than the best current methods, with an accuracy of about 100% and a smaller number of chosen features across five cancer datasets: colon, lung, leukemia, SRBCT, and prostate. The limitation of this study is that the performance evaluation is mainly based on accuracy and feature count.

Hameed et al. [35] introduced a three-phase hybrid method to select and classify high-dimensional microarray data. To achieve this purpose, the author employed Pearson's Correlation Coefficient (PCC) alongside Binary Particle Swarm Optimization (BPSO) or Genetic Algorithm (GA) and numerous classifiers. In the first phase, the methodology utilizes PCC as a filter for feature selection. Subsequently, the second phase involved the application of either BPSO or GA as wrapper methods. The data was classified using five distinct classifiers. The results showed improved classification accuracy, with BPSO outperforming GA in speed and effectiveness across multiple datasets and classifiers. Although the authors compared BPSO with GA, they did not study a broader range of optimization algorithms or hybrid approaches. This highlights the urgent need for a more comprehensive understanding of the best practices in feature selection and classification. As can be observed from the existing works discussed above, the filter methods have been utilized individually [18,24] or combined with the genetic algorithm [23,24,26,27,33,34,35] or wrapper feature selection [29,32] in

order to improve cancer classification on microarray datasets. In contrast, this study proposed integrating Differential Evolution (DE) with some popular filter methods to maximize cancer classification on microarray datasets. DE has several attractive advantages over other competitive optimization algorithms. DE performs well in convergence, although it is straightforward to implement and requires a few parameters to control and low space complexity.

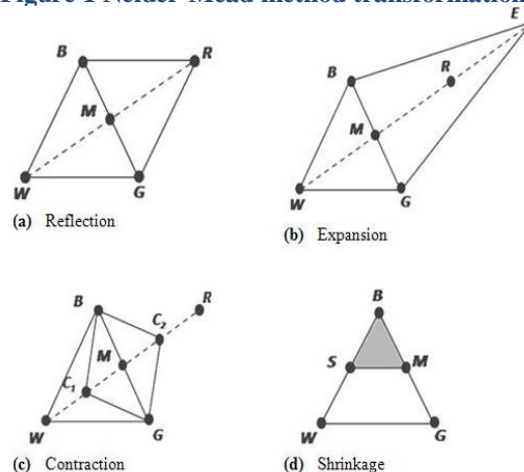
NELDER–MEAD METHOD

The Nelder-Mead algorithm is a direct search method. Thus, it can be viewed as another version of the downhill approach [36]. This is a simple method commonly used in nonlinear optimization technique, which is a well-defined numerical method for problems for which derivatives may not be known. In many numerical tests, the Nelder-Mead method succeeds in obtaining a good reduction in the function value using a relatively small number of function evaluations. Apart from being simple to understand and use, this is the main reason for its popularity in practice. A large subclass of direct search methods, including the Nelder-Mead method, maintain at each step a non-degenerate simplex, a geometric figure in n dimensions of nonzero volume that is the convex hull of $n + 1$ vertices. Each iteration of a simplex-based search begins with $n + 1$ vertices and the associated function values. One or more test points are computed along with their function values, and it continues for a specified number of iterations. In NM procedure the simplex method consists of exactly $n + 1$ solution vectors x_k , where $k = 0, 1, \dots, n$, where n is the number of decision variables or length of the solution vector. Among each of $n + 1$ solutions represent a point in the search space will form a geometrical object in n dimensions called the simplex. Each operation of the Nelder-Mead simplex procedure is performed on the entire solution vector.

The procedure takes an initial simplex as argument and returns the best solution in another final simplex. Every time the solutions in the simplex are ordered in

increasing order by the objective values. The mean m is computed based on the average of n best solutions except the worst solution $n + 1$. The reflection point R of the worst point W is calculated which is shown in Figure 1(a). If the objective value of the reflection point is between the objective value of the best point B and the objective value of the second to worst point G , then only the reflected point is accepted. Otherwise the objective value of the reflection point is less than the objective value of the best point in the simplex a new expansion point E is calculated. The expansion point is illustrated by Figure 1(b). The best of the reflection and the expansion point is accepted to the simplex at the expense of the worst point $n + 1$. If the objective value of the reflection point is larger than the second to worst point n in the simplex a contraction is performed. In case the reflection point has an objective value which is larger than the worst point the contraction point C_1 placed inside the simplex. The inside contraction point is only accepted if the objective is strictly smaller than the worst objective, Otherwise the simplex is shrunk. In case the reflection point has an objective value that is smaller than the objective of the worst point, a contraction point C_2 situated outside the simplex is calculated. The contraction point is only accepted at the expense of the worst point if it has an objective smaller than the reflection point. Otherwise, the simplex is shrunk. The contraction points C_1 and C_2 are shown in Figure 1(c). Only in shrink operation the best point B is kept in a shrink step, the following n points are moved towards to the best point. Figure 1(d) illustrates the shrink point S . The Nelder-Mead procedure returns the solution point from the final simplex which has the smallest objective value. The Pseudo code for Nelder-Mead method is shown in Figure 2. Four scalar parameters must be specified to describe an absolute Nelder-Mead method: coefficients of reflection (α), expansion (χ), contraction (γ), and shrinkage (δ). The general choices used in the standard Nelder-Mead algorithm (Lagarias et al 1998) are $\alpha = 1$, $\chi = 2$, $\gamma = 1$, and $\delta = 1$.

Figure 1 Nelder-Mead method transformations



1. **Order.** Sort the vertices $x_1, x_2, \dots, x_n, x_{n+1}$ of the current simplex so that $f(x_1) \leq f(x_2) \leq \dots \leq f(x_n) \leq f(x_{n+1})$ in ascending order.
2. **Reflect.** Compute the reflection point x_r

$$x_r = m + \alpha (\underline{m} - x_{n+1}) = (1 + \alpha) m - \alpha x_{n+1}$$

$$m = \sum_{i=1}^n \frac{x_i}{n}$$

Where M is the centroid of the n best points (all vertices except for x_{n+1}), evaluate $f_r = f(x_r)$.
 If $f_r \leq f_n$, accept the reflected point x_r and terminate the iteration.
3. **Expand.** If $f_r < f_1$, calculate the expansion point x_e

$$x_e = m + \gamma (\underline{x_r} - m) = (1 + \gamma) m - \gamma x_{n+1}$$
, evaluate $f_e = f(x_e)$.
 If $f_e < f_r$, accept x_e and terminate the iteration; otherwise accept x_r and terminate the iteration.
4. **Contract.** If $f_r \geq f_n$, perform a contraction between m and the better of x_{n+1} and x_r .
 - a. **Outside.** If $f_r \leq f_n < f_{n+1}$, perform an outside contraction:

$$x_{or} = m + \gamma (\underline{x_r} - m) = (1 + \gamma) m - \gamma x_{n+1}$$
, evaluate $f_{or} = f(x_{or})$
 If $f_{or} \leq f_r$, accept x_{or} and terminate the iteration;
 otherwise, go to step 5 (perform a shrink).
 - b. **Inside.** If $f_r \geq f_{n+1}$, perform an inside contraction: calculate

$$x_{ir} = m - \gamma (\underline{m} - x_{n+1}) = (1 - \gamma) m + \gamma x_{n+1}$$
, evaluate $f_{ir} = f(x_{ir})$
 If $f_{ir} < f_{n+1}$, accept x_{ir} and terminate the iteration;
 otherwise, go to step 5 (perform a shrink).
5. **Shrink.** Evaluate f at the n points $v_i = x_i + \delta (x_i - x_1)$, $i = 2, \dots, n+1$.
 The (unordered) vertices of the simplex at the next iteration consist of x_1, v_2, \dots, v_{n+1} .

Figure 2 Pseudo code for one iteration of Nelder-Mead method

DIFFERENTIAL EVOLUTION ALGORITHM

DE is a stochastic, population-based optimization algorithm which is introduced by Storn & Price [37]. DE optimizes a problem by maintaining a population of candidate solutions and creating new candidate solutions by combining existing ones according to its simple formulae, and then keeping whichever candidate solution has the best score or fitness on the optimization problem at hand. In this way the optimization problem is treated as a black box that merely provides a measure of quality given a candidate solution and the gradient is therefore not needed. It is developed to optimize real parameter, real valued functions. Procedure for Differential Evolution is given below.

- Initialize the random solution x_i
 - Calculate the objective function value $f(x_i)$ for all x_i .
 - Select three points $xr1$, $xr2$ and $xr3$ from population and generate perturbed individual using
- o $v_i = xr1 + F \times (xr2 - xr3)$
- Recombine each target vector x_i with perturbed individual generated
 - Calculate the objective function value for u_i .
 - Choose better of the two-function value at target and trial point u_i and x_i for next generation.

The choice of DE parameters F and P_c can have a large impact on optimization performance. F is the real and constant factor which controls the amplification of the differential variation ($xr2 - xr3$). P_c is the crossover constant factor.

Modified Nelder-Mead Method For Biclustering Microarray Gene Expression Data

The NM method minimizes a function of n parameters by comparing the $n + 1$ vertices of a simplex and updating the worst vertex by moving it around a centroid. This simplex method is considered as a fast and simple algorithm. However, when optimizing high dimensional problems, NM method may have poor solution convergence because of that it may not define its moving directions well enough just by simple geometrical movements and it works well for unimodal problem. Therefore, the proposed method MNM considers the median instead of mean and DE is applied before shrinking. The median is the value that divides the distribution exactly into halves. The median is the balance point of the distribution. The main advantage of the median is that it is not affected by outliers as the mean. For skewed data the median provides a better estimate than mean. For outlier data the mean can be a misleading measure of central tendency and the median

value or the mode value are typically more accurate measures. DE uses floating-point for encoding population members and arithmetic operations for mutation. DE finds the true global minimum regardless of the initial parameter values and it converges quickly. For these reasons DE is combined with Nelder-Mead for finding the global optima. In DE, closer the population gets to the global optimum, more the distribution will shrink and therefore reinforce the generation of smaller difference vectors.

Vertex Representation

Each vertex is represented as a candidate solution for the problem. Solutions are encoded by means of binary strings of length $N+M$, where N and M are the number of rows and columns of the expression data respectively Mitra & Banka [38]. A bit is set to one if the corresponding gene or condition are present in the bicluster, and reset to zero otherwise. So the individual dimension of solution is represented by a real number. Figure 4 shows the representation of solution and its mapped bicluster representation. The mapping function of solution into a binary string representation of a bicluster is given in Equation (1) as follows:

$$y_{ij} = \begin{cases} x_{ij} \geq 0.5 & 1 \\ \text{otherwise} & 0 \end{cases}$$

(1)

Where

x_{ij} - Random value generated for j th gene/condition of i th point

y_{ij} - Binary string representation of bicluster of x_{ij}

In y_{ij} , if a bit is set to 1 then the corresponding gene or condition belongs to the encoded bicluster; otherwise, it is not.

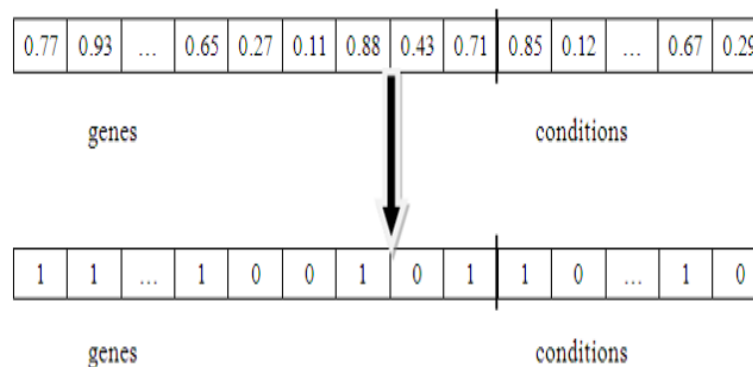


Figure 3. Representation of vertex and its mapping to biclusters

In NM and MNM based biclustering method, the fitness function of an individual is determined by evaluating the MSR and row variance. The fitness value of each vertex is calculated. Pseudo code for one iteration of Modified Nelder-Mead is shown in Figure 4.

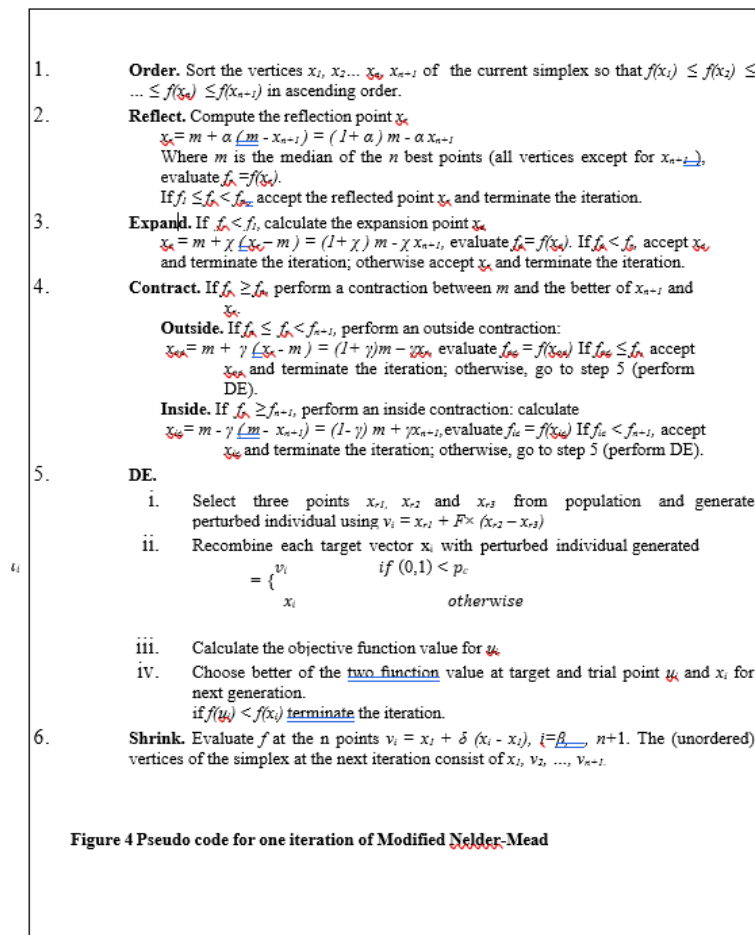


Figure 4 Pseudo code for one iteration of Modified Nelder-Mead

Experimental Results and Analysis

Experimental Setup

The Nelder-Mead and Modified Nelder-Mead algorithms presented for the bicluster problem are implemented in MATLAB and run on an Intel i3 3.7 GHz. The minimum fitness value is obtained for 20 biclusters with the stopping criterion is up to the maximum iteration 1000. A good choice for F and P_c are 0.5 and 0.8 respectively. Table 1 shows the parameter and its value used in this work.

Table 1 Parameter and its value

Parameter	Value
Coefficient of reflection (α)	1
Coefficient of expansion (χ)	2
Coefficient of contraction (γ)	0.5
Coefficient of shrinkage (δ)	0.5
Constant factor (F)	0.5
Crossover constant (P_c)	0.8
Number of biclusters	20
Number of Iterations	1000

Bicluster Extraction for Yeast Cell Cycle and Human B-Cell Lymphoma Expression Dataset

Granting to the problem formulation of an extracted bicluster should be satisfying a homogeneity criterion. The bicluster should satisfy two requirements simultaneously. The expression levels of each gene within the bicluster should be similar over the range of conditions. It means that it should have a low MSR score. On the other hand, the bicluster row variance should be high. The MSR represents the variance of the selected genes and conditions with respect to the homogeneity of the bicluster and row variance removes the simple bicluster. To quantify the biclusters homogeneity and size satisfy the Coherence Index (CI) is used as a measure for evaluating their goodness Mitra & Banka [38]. CI is defined as the ratio of MSR score to the size of the formed bicluster. The size of a bicluster increases while CI proportionately decreases. Table 2 and 3 show the experimental results obtained for yeast cell cycle data and human lymphoma data respectively. Totally five biclusters are chosen randomly from the total number of biclusters. Figure 5 and 6 show the fitness value obtained for yeast cell cycle data and human lymphoma data respectively. For both data sets the proposed MNM

outperforms NM algorithm because the DE allows Nelder-Mead to escape from local optimum and successfully continue to the global optimum.

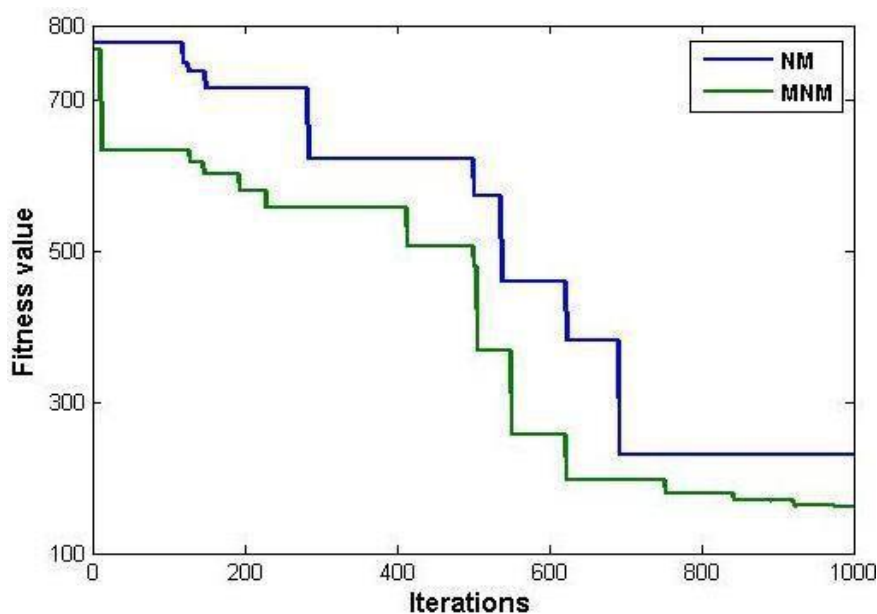


Figure 5. Plot of number of iterations versus fitness value on yeast cell cycle data

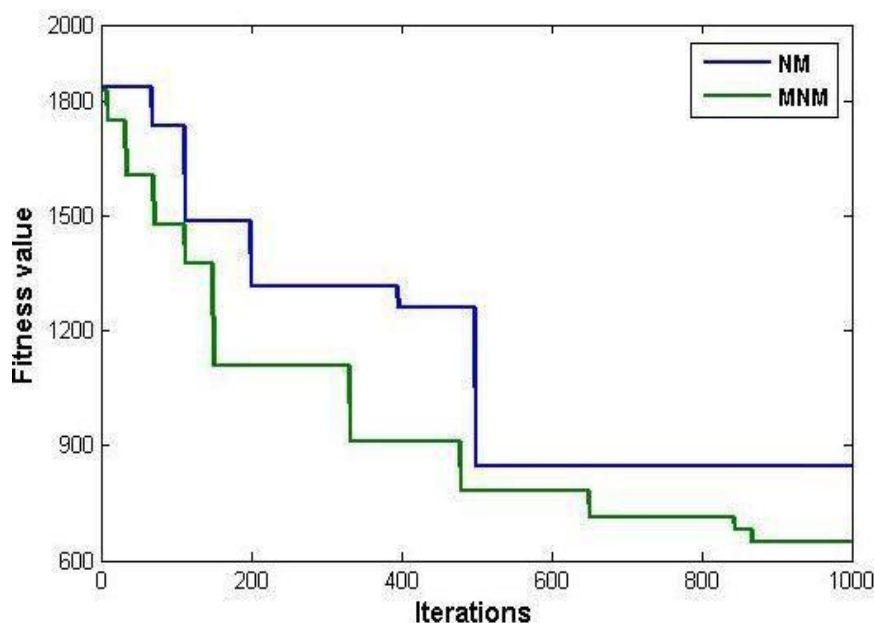


Figure 6: Plot of number of iterations versus fitness value on human B-cell lymphoma data

Table 2 summarizes the best biclusters for yeast cell cycle expression data after 1000 generations. The largest sized bicluster is found at MSR=212.22, with coherence index CI being minimal and indicating the goodness of the discovered partitions. The minimum value of CI is 0.0505 with a corresponding size of 4200 being the best in the table. As mentioned earlier, a low mean squared residue indicates a high coherence of the discovered biclusters.

Table 2. Extracted biclusters for yeast cell cycle data

Bicluster	Genes	Conditions	Volume	MSR	Row Variance	CI
BC2	155	10	1550	156.79	812.89	0.1011
BC8	365	7	2555	171.91	695.37	0.0672
BC4	406	6	2436	179.11	715.74	0.0735
BC15	356	9	3204	184.75	823.42	0.0576
BC4	420	10	4200	212.22	912.41	0.0505

Table 3 summarizes the best biclusters for Human B-cell data after 1000 generations. The largest sized bicluster is found at MSR=856.93, with coherence index CI being minimal and indicating the goodness of the discovered partitions. The

minimum value of CI is 0.0923 with a corresponding size of 9275 being the best in the table. As mentioned earlier, a low mean squared residue indicates a high coherence of the discovered biclusters.

Table 3. Extracted biclusters for human B-cell lymphoma data

Bicluster	Genes	Conditions	Volume	MSR	Row Variance	CI
BC1	295	25	7375	756.45	2272.20	0.1025
BC5	245	33	8085	782.67	2129.33	0.0968
BC6	302	28	8456	810.20	2321.52	0.0958
BC14	273	32	8736	825.11	2385.19	0.0944
BC9	265	35	9275	856.93	2479.76	0.0923

Figure 7 depicts the gene expression profile of the largest bicluster, corresponding to MSR=212.22. The gene expression values in the range 150 to 350 indicate the highly dense profiles of the co regulated genes having little or no fluctuations under the selected conditions of the bicluster. It has the highest row variance is 912.41 whereas the MSR is 212.22. In terms of fitness, this is the most “interesting” bicluster which has largest volume 4200 with the lowest MSR. Moreover, MNM tries to find highly row-variant biclusters instead of trivial biclusters.

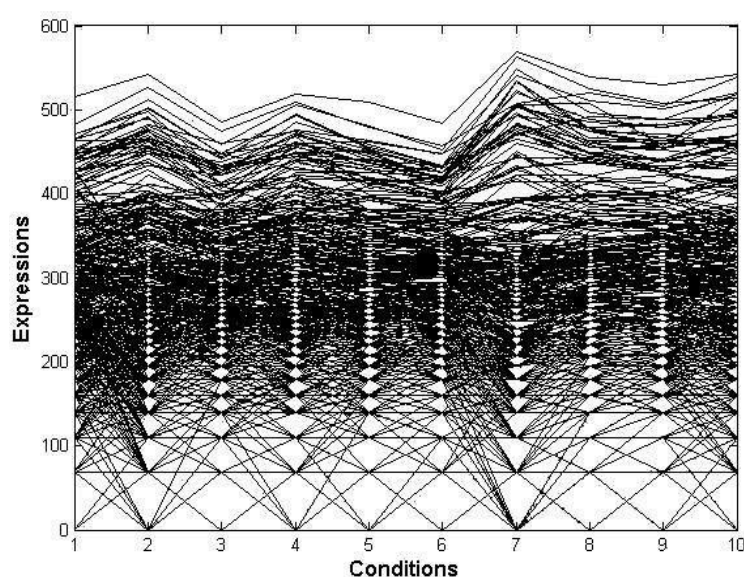


Figure 7: Gene expression profile of the largest bicluster on yeast cell cycle data

Figure 8 depicts the gene expression profile of the largest bicluster, corresponding to MSR=856.93. It has the highest row variance of 2479.76, whereas the MSR is 856.93. In terms of fitness value 856.93, this is the most “interesting” bicluster which has the largest volume 9275 with the lowest MSR. The gene expression values in the range -100 to 100 indicate the highly dense profiles of the co regulated genes having little or no fluctuations under the selected conditions of the bicluster. However, there also exist a few genes having large expression values. Perhaps, this is because of the presence of a large number of missing values (12.3%) that are replaced by random numbers between -800 and 800, some of which remain in the biclusters without violating the homogeneity restriction. Sometimes this can also occur when a few genes have large variation in their expression values get included while continuing to satisfy the homogeneity constraint of the bicluster.

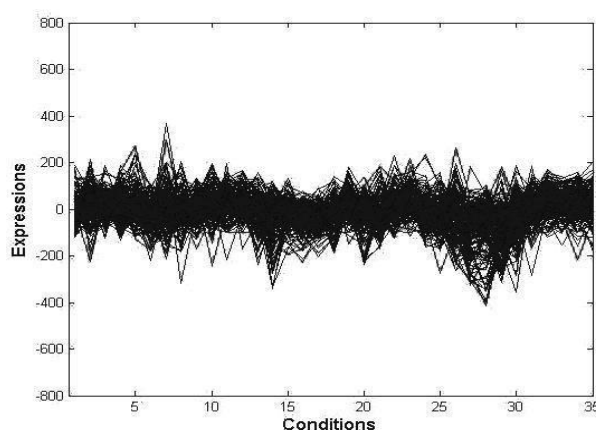


Figure 8: Gene expression profile of the largest bicluster on human B-cell lymphoma data

Comparative Analysis Based on MSR

The Table 4 depicts the results of the proposed method that is compared with the well-known existing methods namely FLOC [39], CC[40], SEBI [41], SMOB [42] and PCOBA [43] on yeast cell cycle expression dataset. The high value of MSR shows that the bicluster is weakly coherent while a low value MSR indicates that it is highly coherent [40]. FLOC and PCOBA often failed to find a homogenous block structure that uses a probabilistic approach to find biclusters. Similarly, CC algorithm gives a limited size of biclusters for large MSR. FLOC is able to locate large biclusters for minimum MSR compared with CC. However, extracted bicluster is not significant. SEBI average volume is 209.92 for MSR of 205.18. The bicluster found by SMOB is interesting; however, this is extract very small size of bicluster average MSR of 206.17. Next NM method returns the largest bicluster, nevertheless average MSR of NM is larger than all the other methods. In the case of MNM, average MSR is better than that of all other algorithms. Even so average volume is far better than that of SEBI and SMOB algorithm. Although FLOC shows better residue scores than CC, SEBI, SMOB, PCOBA and NM did, they were not superior to MNM.

Table 4: Comparative analysis on yeast cell cycle data

Method	Average MSR	Average volume	Average no. of genes	Average no. of conditions
FLOC	187.44	1825.78	195.00	12.20
CC	204.29	1576.98	167.00	12.00
SEBI	205.18	209.92	13.61	15.25
SMOB	206.17	453.48	27.28	15.46
PCOBA	219.15	1321.30	92.40	14.30
NM	234.25	2876.36	312.11	7.30
MNM	180.56	2903.32	307.67	8.04

Table 5 gives performance comparison of MNM for Human B-cell Lymphoma dataset with that of CC, SEBI and the algorithm SMOB. In this dataset the average number of MSR and average volume of the biclusters obtained are better than the other algorithms. Average number of conditions is greater than CC and NM. However, CC algorithm is capable of finding biclusters characterized by a higher volume with minimum MSR than the ones found by SEBI and SMOB. In the case of MNM algorithm, average number of genes is better than all the other methods. Therefore, the proposed MNM method generates good quality of biclusters with comparatively smaller residue values. It frequently gives significant improvements in the first few iterations and produces quite satisfactory results.

Table 5: Comparative analysis on human B-cell lymphoma data

Method	Average MSR	Average volume	Average no. of genes	Average no. of conditions
CC	850.04	4595.98	269.22	24.5
SEBI	1028.84	615.84	14.07	43.57
SMOB	1019.16	709.13	11.60	78.47
NM	912.21	7918.28	256.41	28.50
MNM	832.09	8226.55	284.20	30.11

Statistical Relevance

With the intention of evaluating the statistical relevance of MNM algorithm, the results of the proposed method are compared with CC, ISA, Bimax, OPSM and BiMine on yeast cell cycle expression data from [44] by using web-tool of FuncAssociate (Roth lab 2008). The FuncAssociate computes the adjusted significance scores for each biclusters. In fact, the adjusted significance scores assess genes in each bicluster by computing adjusted p-values, which indicates how well they match with the different GO categories. Indeed the biclusters that have an adjust p-value lower than the 5% are considered as overrepresented. This means that majority of genes of a bicluster have common biological characteristics. Figure 9 represents the different values of significant scores p-value for each algorithm over the percentage of total extracted biclusters. Analysis shows that the 100% of the tested biclusters under BiMine, OPSM, Bimax and MNM have p-value 5% and 1%. Finally, 65% of extracted biclusters with MNM have p-value = 0.001%, while those of NM, BiMine, OPSM, Bimax, ISA and CC have 47%, 51%, 22%, 64%, 32%, 10% .Note that MNM performs well for all p-values compared to other techniques. Also, MNM performs well for all cases of p-value (p-value = 5%, p-value = 1%, p-value = 0.5%, p-value = 0.1% and p-value = 0.001%).

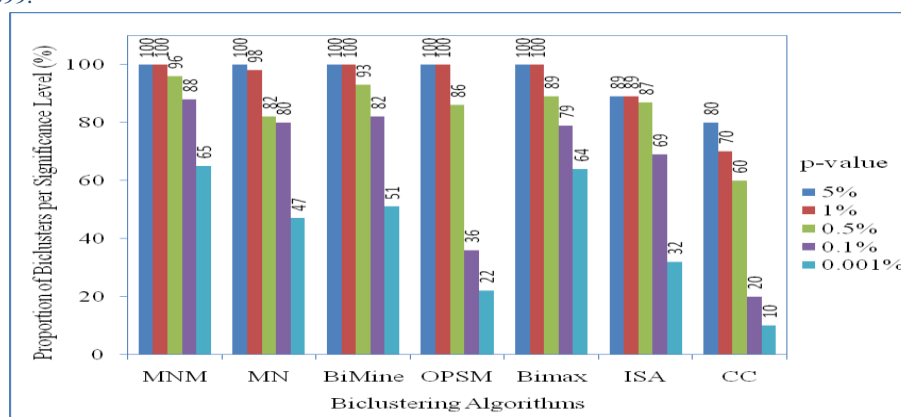


Figure 9 Proportions of biclusters significantly enriched by GO annotations on Yeast cell cycle data

Biological Annotation for Yeast Cell Cycle Using GO Term Finder Toolbox

GO Term Finder is a tool available in the Saccharomyces Genome Database (SGD), in an attempt to identify the biological annotations for the biclusters (Stanford University 2004). It is designed to search for the significant shared GO terms of the groups of genes and provides users with the means to identify the characteristics that the genes may have in common. Table 6 lists the significant shared GO terms or parent of GO terms used to describe the set of genes in each bicluster for the process, function and component ontologies. For example, to the bicluster BC2, the genes are mainly involved in translation process, structural constituent of ribosome activity and cytosolic ribosome component. The tuple (n=69, p=2.26×10⁻⁷²) represents that out of 155 genes in bicluster BC2, 69 genes belong to in cytoplasmic translation process, and the statistical significance is given by the p-value of p=2.26×10⁻⁷². Next, the tuple (n=68, p=7.82×10⁻⁶²) represents that out of 155 genes in bicluster BC2, 68 genes belong to structural constituent of ribosome activity function, and the statistical significance is given by the p-value of p=7.82×10⁻⁶². Finally, the genes 69 out of 155 belong to component of cytosolic ribosome and the corresponding p-value is p=3.73×10⁻⁷².

Table 6: Significant GO terms of yeast cell cycle data

Process	Function	Component
Cytoplasmic translation (n=69, p=2.26×10 ⁻⁷²)	Structural constituent of ribosome (n=68, p=7.82×10 ⁻⁶²)	Cytosolic ribosome (n=69, p=3.73×10 ⁻⁷²)
Ribosome biogenesis (n=71, p=1.82×10 ⁻⁴⁵)	Structural molecule (n=68, p=5.85×10 ⁻⁴⁸)	Cytosolic part (n=69, p=3.93×10 ⁻⁶²)
Cellular metabolic process (n=144, p=1.41×10 ⁻²⁴)	RNA binding (n=46, p=9.66×10 ⁻⁸)	Organelle part (n=117, p=4.85×10 ⁻²¹)

Figure 10, depicts the significant GO terms or parents of GO terms for a set of 20 genes along with their p-values, with the significance being indicated in terms of the different colors displayed. It shows the branching of a generalized molecular function into sub-functions like structural molecule activity and protein tag etc., which are then clustered gene-wise to produce the final result. Moreover out of 20 genes, the 9 genes (RPS21A, RPL40B, RPL8B, RPL15A, RPS0B, RPL22A, RPL10, RPS31, RPL37A) are involved in structural constituent of ribosome. Further the corresponding p-value is very small (p= 6.98×10⁻⁰⁸) which shows that there is very less probability to obtain the gene cluster in random. Those result means that the proposed MNM biclustering approach can find biologically meaningful biclusters.

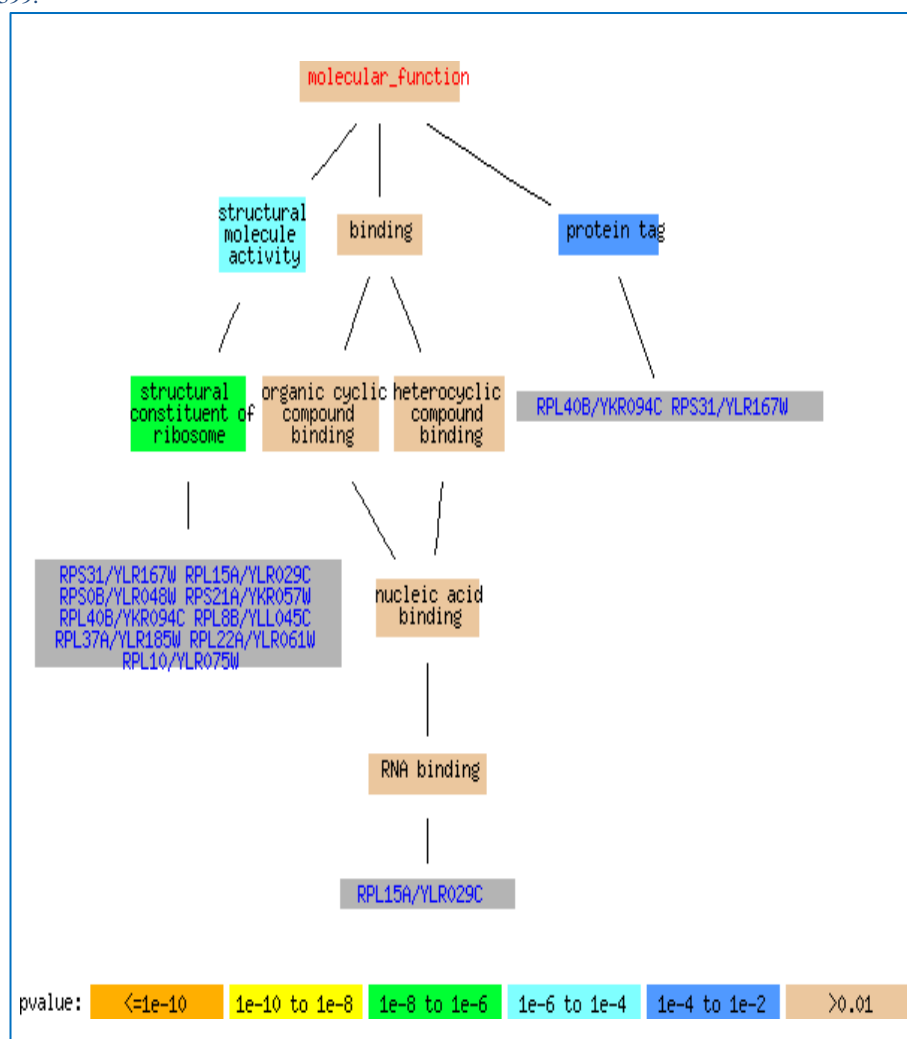


Figure 10 Gene Ontology biological functions of yeast cell cycle data (20 genes)

SUMMARY

The Modified Nelder-Mead algorithm for biclustering microarray gene expression data is proposed to overcome the poor convergence problem of NM method. It focuses on finding coherent biclusters with lower MSR and higher row variance. In Nelder-Mead method the median is measured instead of mean. The median provides much better estimates in place of mean. Before shrinking operation the differential evolution is applied to obtain global minimal solution. A qualitative measure of the formed biclusters with a comparative assessment of results are provided on two benchmark gene expression datasets to demonstrate the effectiveness of the proposed method. Biological validation of the selected genes within the biclusters is provided by publicly available GO consortium. The patterns present a significant biological relevance in terms of related biological processes, components and molecular functions in a species-independent manner. In conclusion, it is found that the Modified Nelder-Mead approach gives a better result over the conventional Nelder-Mead method and existing biclustering algorithms.

REFERENCES

1. Kourou, K.; Exarchos, T.P.; Exarchos, K.P.; Karamouzis, M.V.; Fotiadis, D.I. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.* 2014, 13, 8–17. [Google Scholar] [CrossRef] [PubMed]
2. Yu, K.H.; Lee, T.L.M.; Yen, M.H.; Kou, S.C.; Rosen, B.; Chiang, J.H.; Kohane, I.S. Reproducible machine learning methods for lung 451 cancer detection using computed tomography images: Algorithm development and validation. *J. Med. Internet Res.* 2020, 22, 16709. [Google Scholar] [CrossRef] [PubMed]
3. Felman, A. What to Know about Breast Cancer. *Medical News Today*. Available online: <https://www.medicalnewstoday.com/articles/37136> (accessed on 20 December 2023).
4. Malebary, S.J.; Hashmi, A. Automated Breast Mass Classification System Using Deep Learning and Ensemble Learning in Digital Mammogram. *IEEE Access* 2021, 9, 55312–55328. [Google Scholar] [CrossRef]
5. Zahoor, M.M.; Qureshi, S.A.; Bibi, S.; Khan, S.H.; Khan, A.; Ghafoor, U.; Bhutta, M.R. A New Deep Hybrid Boosted and Ensemble

- Learning-Based Brain Tumor Analysis Using MRI. *Sensors* 2022, 22, 2726. [Google Scholar] [CrossRef]
6. Hashmi, A.; Barukab, O. Dementia Classification Using Deep Reinforcement Learning for Early Diagnosis. *Appl. Sci.* 2023, 13, 1464. [Google Scholar] [CrossRef]
7. Hashmi, A.; Osman, A.H. Brain Tumor Classification Using Conditional Segmentation with Residual Network and Attention Approach by Extreme Gradient Boost. *Appl. Sci.* 2022, 12, 10791. [Google Scholar] [CrossRef]
8. Ostrom, Q.T.; Cioffi, G.; Gittleman, H.; Patil, N.; Waite, K.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012–2016. *Neuro Oncol.* 2019, 21, v1–v100. [Google Scholar] [CrossRef] [PubMed] [PubMed Central]
9. Musheer, R.A.; Verma, C.K.; Srivastava, N. Novel machine learning approach for classification of high-dimensional microarray data. *Soft Comput.* 2019, 23, 13409–13421. [Google Scholar] [CrossRef]
10. Singh, R.K.; Sivabalakrishnan, M. Feature Selection of Gene Expression Data for Cancer Classification: A Review. *Procedia Comput. Sci.* 2015, 50, 52–57. [Google Scholar] [CrossRef]
11. Wang, L. Feature selection in bioinformatics. In *Independent Component Analyses, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering X*; SPIE: Baltimore, MD, USA, 2012; Volume 8401, Available online: <https://hdl.handle.net/10356/84511> (accessed on 20 December 2023).
12. Song, Q.; Ni, J.; Wang, G. A Fast Clustering-Based Feature Subset Selection Algorithm for High-Dimensional Data. *IEEE Trans. Knowl. Data Eng.* 2013, 25, 1–14. [Google Scholar] [CrossRef]
13. Saeys, Y.; Inza, I.; Larrañaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* 2007, 23, 2507–2517. [Google Scholar] [CrossRef] [PubMed]
14. Wang, A.; Liu, H.; Yang, J.; Chen, G. Ensemble feature selection for stable biomarker identification and cancer classification from microarray expression data. *Comput. Biol. Med.* 2022, 142, 105208. [Google Scholar] [CrossRef] [PubMed]
15. De Souza, J.T.; De Francisco, A.C.; De Macedo, D.C. Dimensionality Reduction in Gene Expression Data Sets. *IEEE Access* 2019, 7, 61136–61144. [Google Scholar] [CrossRef]
16. Bhui, N. Ensemble of Deep Learning Approach for the Feature Selection from High-Dimensional Microarray Data; Springer: Berlin/Heidelberg, Germany, 2022; pp. 591–600. [Google Scholar] [CrossRef]
17. Alhenawi, E.; Al-Sayyed, R.; Hudaib, A.; Mirjalili, S. Feature selection methods on gene expression microarray data for cancer classification: A systematic review. *Comput. Biol. Med.* 2022, 140, 105051. [Google Scholar] [CrossRef]
18. Abdulla, M.; Khasawneh, M.T. G-Forest: An ensemble method for cost-sensitive feature selection in gene expression microarrays. *Artif. Intell. Med.* 2020, 108, 101941. [Google Scholar] [CrossRef]
19. Foster, K.R.; Koprowski, R.; Skufca, J.D. Machine learning, medical diagnosis, and biomedical engineering research—Commentary. *Biomed. Eng. OnLine* 2014, 13, 94. [Google Scholar] [CrossRef]
20. MS, K.; Rajaguru, H.; Nair, A.R. Enhancement of Classifier Performance with Adam and RanAdam Hyper-Parameter Tuning for Lung Cancer Detection from Microarray Data—In Pursuit of Precision. *Bioengineering* 2024, 11, 314. [Google Scholar] [CrossRef]
21. Elbashir, M.K.; Almotilag, A.; Mahmood, M.A.; Mohammed, M. Enhancing Non-Small Cell Lung Cancer Survival Prediction through Multi-Omics Integration Using Graph Attention Network. *Diagnostics* 2024, 14, 2178. [Google Scholar] [CrossRef]
22. Zamri, N.A.; Aziz, N.A.A.; Bhuvaneswari, T.; Aziz, N.H.A.; Ghazali, A.K. Feature Selection of Microarray Data Using Simulated Kalman Filter with Mutation. *Processes* 2023, 11, 2409. [Google Scholar] [CrossRef]
23. Ali, W.; Saeed, F. Hybrid Filter and Genetic Algorithm-Based Feature Selection for Improving Cancer Classification in High-Dimensional Microarray Data. *Processes* 2023, 11, 562. [Google Scholar] [CrossRef]
24. Elemam, T.; Elshrkawey, M. A Highly Discriminative Hybrid Feature Selection Algorithm for Cancer Diagnosis. *Sci. World J.* 2022, 2022, 1056490. [Google Scholar] [CrossRef] [PubMed]
25. Abasabadi, S.; Nematzadeh, H.; Motameni, H.; Akbari, E. Hybrid feature selection based on SLI and genetic algorithm for microarray datasets. *J. Supercomput.* 2022, 78, 19725–19753. [Google Scholar] [CrossRef] [PubMed]
26. Saeed, F.; Almutiri, T. A Hybrid Feature Selection Method Combining Gini Index and Support Vector Machine with Recursive Feature Elimination for Gene Expression Classification. *Int. J. Data Min. Model. Manag.* 2022, 14, 41–62. [Google Scholar] [CrossRef]

27. Xie, W.; Fang, Y.; Yu, K.; Min, X.; Li, W. MFRAG: Multi-Fitness RankAggreg Genetic Algorithm for biomarker selection from microarray data. *Chemom. Intell. Lab. Syst.* 2022, 226, 104573. [Google Scholar] [CrossRef]
28. Dash, R. An Adaptive Harmony Search Approach for Gene Selection and Classification of High Dimensional Medical Data. *J. King Saud Univ.-Comput. Inf. Sci.* 2021, 33, 195–207. [Google Scholar] [CrossRef]
29. Almutiri, T.; Saeed, F.; Alassaf, M.; Hezzam, E.A. A Fusion-Based Feature Selection Framework for Microarray Data Classification. In *Proceedings of the International Conference of Reliable Information and Communication Technology*, Online, 22–23 December 2021; pp. 565–576. [Google Scholar] [CrossRef]
30. Kilicarslan, S.; Adem, K.; Celik, M. Diagnosis and classification of cancer using hybrid model based on ReliefF and convolutional neural network. *Med. Hypotheses* 2020, 137, 109577. [Google Scholar] [CrossRef] [PubMed]
31. Baliarsingh, S.K.; Vipsita, S.; Dash, B. A new optimal gene selection approach for cancer classification using enhanced Jaya-based forest optimization algorithm. *Neural Comput. Appl.* 2020, 32, 8599–8616. [Google Scholar] [CrossRef]
32. Almugren, N.; Alshamlan, H. A Survey on Hybrid Feature Selection Methods in Microarray Gene Expression Data for Cancer Classification. *IEEE Access* 2019, 7, 78533–78548. [Google Scholar] [CrossRef]
33. Sayed, S.; Nassef, M.; Badr, A.; Farag, I. A Nested Genetic Algorithm for feature selection in high-dimensional cancer Microarray datasets. *Expert Syst. Appl.* 2019, 121, 233–243. [Google Scholar] [CrossRef]
34. Ghosh, M.; Adhikary, S.; Ghosh, K.K.; Sardar, A.; Begum, S.; Sarkar, R. Genetic algorithm based cancerous gene identification from microarray data using ensemble of filter methods. *Med. Biol. Eng. Comput.* 2019, 57, 159–176. [Google Scholar] [CrossRef]
35. Hameed, S.S.; Muhammad, F.F.; Hassan, R.; Saeed, F. Gene Selection and Classification in Microarray Datasets using a Hybrid Approach of PCC-BPSO/GA with Multi Classifiers. *J. Comput. Sci.* 2018, 14, 868–880. [Google Scholar] [CrossRef]
36. Nelder, JA & Mead, R 1965, 'A simplex method for function minimization', *Computer Journal*, vol. 7, no. 5, pp. 308-313.
37. Storn, R & Price, K 1997, 'Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces', *Journal of Global Optimization*, vol. 11, no. 4, pp. 341 - 359.
38. Mitra, S & Banka, H 2006 'Multi-objective evolutionary biclustering of gene expression data', *Pattern Recognition*, vol. 39, no. 12, pp. 2464 - 2477.
39. Yang, J, Wang, H, Wang, W & Yu, P 2003, 'Enhanced biclustering on expression data' : proceedings of the Third IEEE Symposium on BioInformatics and BioEngineering, pp. 321-327.
40. Cheng, Y & Church, GM 2000, 'Biclustering of expression data' : proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology, pp. 93-103.
41. Divina, F & Aguilar-Ruiz, JS 2006, 'Biclustering of expression data with evolutionary computation', *IEEE Transactions Knowledge and Data Engineering*, vol. 18, no. 5, pp. 590-602.
42. Divina, F & Aguilar-Ruiz, JS 2007, 'A multi-objective approach to discover biclusters in microarray data': proceedings of the ninth annual conference on Genetic and evolutionary computation, pp. 385-392.
43. Joung, JG, Kim, SJ, Shin, SY & Zhang, BT 2012, 'A probabilistic coevolutionary biclustering algorithm for discovering coherent patterns in gene expression dataset', *BMC Bioinformatics*, vol. 13, no. 1, p. S12.
44. Ayadi, W & Hao, JK 2014, 'A Memetic Algorithm for Discovering Negative Correlation Biclusters of DNA Microarray Data', *Neurocomputing*, vol. 145, no. 7, pp. 14-22.