# Gene Selection and Classification Using Quantum Moth Flame Optimization Algorithm

1. Ali Dabba[a,c,d]
2. Abdelkamel Tari[a,e]
3. Samy Meftali[b,d]

[a]*Faculty of Sciences,Computer Science Department, Abderrahmane Mira University,Bejaia,Algeria*
[b]*University of Lille, France*
[c]*Faculty of Mathematics and Computer Science, Computer Science Department, Mohamed Boudiaf University, M'sila, Algeria.*
[d]*Research center in Computer Science, Signal and Automatic Control of Lille – CRIStAL*
[e]*Laboratory of Medical Computing - LIMD*

***ABSTRACT* - In this paper, we present a new swarm intelligence algorithm for gene selection called quantum moth flame optimization algorithm (QMFOA), which based on hybridization between quantum computation and moth flame optimization algorithm (MFOA). The purpose of QMFOA is to identify a small gene subset that can be used to classify samples with high accuracy. The QM-FOA has a simple two-phase approach, the first phase is a pre-processing that uses to address the difficulty of high-dimensional data, which measure the redundancy and the relevance of the gene, in order to obtain the relevant gene set. The second phase is hybridization among MFOA, quantum computing, and support vector machine (SVM) with leave-one-out cross-validation (LOOCV), in order to solve the gene selection problem. The main objective of the second phase is to determine the best relevant gene subset of all genes obtained in the first phase.**

**In order to assess the performance of the proposed QMFOA, we test it on six Microarray datasets. Experimental results show that QMFOA provides great classification accuracy in comparison to some known algorithms.**

***KEYWORDS* - *Genes expression, Feature Selection, Moth Flame Optimization, Algorithm Quantum Computing, Microarray Data, Cancer Classification, Bio-inspired, Algorithms Molecular, Biology Optimization, Algorithms, Evolutionary, Algorithms Swarm, Intelligence.***

## 1. INTRODUCTION

Gene selection is a branch of feature selection, which establishes an evident approach to reducing dimensionality and over-fitting [17]. The main task of gene selection is to find the best subset of genes from all possible choices by filtering out irrelevant, redundant and noisy genes [25]. To achieve good classification accuracy, it is important to choose the most pertinent genes that are necessary and sufficient to describe the target concept, like gives some aspects of functional genomics. In addition, to find an optimal small set of relevant genes has been proven to be an NP-complete problem [5, 7].

In literature, several gene selection methods have been proposed and can be organized into three categories including filter, wrapper, and embedded methods [18, 9]. Filter methods utilize essentially the general statistical properties of the training data at hand without using any learning algorithm. Although these methods are fast but have rather poor performance. In contrast, the wrapper methods select a set of discriminatory features by using a predetermined learning algorithm. The interest of these methods is that the chosen subset is perfectly adapted to the classifier. How- ever, the wrapper methods are more costly in computational time because each evaluation of a feature subset requires a training model, in which the computational complexity depends on the complexity of the learning model used [8]. Embedded methods are similar to wrapper approaches by combining the exploration process with a learning algorithm [10], which are an extension of wrapper approaches and undertake feature selection in the process of classifier training. The advantage of these methods is that the classifier provides important information that guides the search, which makes these methods more efficient than wrapper methods.

In recent years, quantum computing has been proposed in the literature [19, 20] as a more effective technique than classical computing. In other words, quantum physics has been used to build a new kind of computers, called quantum computers [20]. Unlike classical computers that deal with binary digits (bits), the basic unit a quantum bit (Q-bit), in addition the usual f0f and f1f states, a Q-bit can also in any superposition of these two states [13]. Therefore, the best suggestion right now is to use quantum algorithms and apply them to classical computers.

Over the past few years, many algorithms have been proposed to solve gene selection using quantum fields (quantum computing). Among them, Cluster QGA have been proposed by [23], which uses clustering to choose a small set of non-redundant representative genes and then applies the Quantum Genetic Algorithm to define a minimal set of non-redundant and relevant genes. The authors in [26] have proposed an approach called binary quantum-behaved particle swarm optimization (BQPSO). This approach coupling between PSO, quantum computing, and support vector machine (SVM) with leave-one-out cross-validation to solve gene selection, which is a discretized version of the original QPSO for binary 0-1 optimization problems. In addition, the $GQA_{SY\ M}$ [1] has been proposed as a hybrid approach between the Genetic Quantum Algorithm and the Support Vector Machines classifier to gene selection and classification of Microarray Data. The main goal of this algorithm is to identify a small subset of genes that could be used to separate two classes of samples with high accuracy.

In this work, we propose a new algorithm called Quantum Moth Flame Optimization Algorithm (QMFOA), in order to find the best gene subset to provide high classification accuracy to cancer Microarray data. The QMFOA inherits parallelism, decentralization, and cooperation of swarm intelligence algorithm (bio-inspired algorithm, specifically MFOA) to solve the gene selection problem. For solve this problem, the QMFOA uses a hybrid model that uses several techniques: Quantum field (Quantum computing), Moth Flame Optimization Algorithm (MFOA), Mini- mum Redundancy-Maximum Relevance (mRMR), and a Support Vector Machine (SVM) with Leave One Out Cross Validation (LOOCV).

In order to prove the advantages of proposed QMFOA, we have tested QMFOA on six well-known datasets issued of Microarray experiments treating cancer and compared our results with several recently published algorithms in the literature. The experimental results have shown that QMFOA can achieve better performance of classification accuracy with a competitive number of genes selected i.e., it is able to provide a minimum number of genes to obtain the highest classification accuracy for solving the gene selection problem in both binary and multi classes.

The remainder of the paper is organized as follows: Section 2 presents our proposed QMFOA approach to gene selection. The experimental results and discussions are included in Section 3. Finally, the conclusion is given in Section 4.

## 2. THE PROPOSED ALGORITHM FOR THE GENE SELECTION PROBLEM

In this section, we propose a new algorithm called Quantum Moth Flame Optimization Algorithm (QMFOA) for gene selection and classification of high dimensional Microarray data. This work is based on a hyridizing Moth- Flame Optimization Algorithm (MFOA) with concepts resulting from the quantum field to provide solutions for a gene selection problem. However, the QMFOA purpose is to select small samples of informative genes amongst thousands of them.

The principle of the proposed algorithm consists of a two-phase approach. In the first phase, instead of to use the full set of available genes, we use preprocessing to select a relatively smaller set of non-redundant and relevant genes and that passed on to the second phase of QMFOA for the effective selection of a minimal set of informative genes. In order to guarantee this, we start by normalization of Microarray data with the Min-Max method that can guarantee a stable convergence of weights and biases [14]. Furthermore, we use the statistical technique of Minimum Redundancy-Maximum Relevance (mRMR) to measure the relevance and redundancy of selected genes, in order to reduce the high number of genes by eliminating genes redundancy [21].

The second phase of QMFOA is applied to the set of *d* representative genes that were obtained in the first phase of QMFOA in order to find a minimal subset of the relevant genes (.i.e, the maximum number of genes in a moth (individual)). Like any bio-inspired algorithms, this algorithm is based on a population of solutions that is preserved through several generations, which seeks the best-fitted solution to the gene selection problem, evaluating the gene subset of those included in the current population. The fundamental idea of this phase is to combine the MFOA and quantum fields with the SVM classifier. The fitness function uses the SVM classifier with the Leave One Out Cross Validation (LOOCV) method and the percentage of genes that are not selected, which is applied in order to evaluate and validate the provided solutions. The main goal of QMFOA is to select a high accuracy genes subset that includes a smaller number of genes. Finally, the better gene subset obtained by QMFOA will be evaluated using the SVM classifier.

### 2.1 Representation of Candidate Solutions

For QMFOA, the moth (individual) represents a gene subset of the *dsubset*: set of genes; obtain maximum number *d* of genes in an individual.
The moth population containing *n* and the number *d* of Q-bits. it is represented as $QM = \{Qm_1, Qm_2, ....., Qm_d\}$,

Where, $Qm_i$ ($i$ = 1, 2,..., $n$) is the $i^{th}$ moth. Each $Qm_i$ quantum moth represents as follows (Eq.1):

$$Qm_1 = \begin{vmatrix} \cos(\theta_{i,1}) & \cos(\theta_{i,2}) & \cos(\theta_{i,3}) & \dots\dots & \cos(\theta_{i,d}) \\ \sin(\theta_{i,1}) & \sin(\theta_{i,2}) & \sin(\theta_{i,3}) & \dots\dots & \sin(\theta_{i,d}) \end{vmatrix}$$

Where, $d$ is the number of Q-bits used in each quantum moth's representation, $\theta_{i,k}$ ($k$ = 1, ....., $d$) with $\theta_{i,k} \epsilon [0, \frac{\pi}{2}]$ represents rotation angle and satisfy the normalization condition
$|\cos(\theta_{i,k})|^2 + |\sin(\theta_{i,k})|^2 = 1$ with

$|cos(\theta_{i,k})|^2$ *the probability of rejecting* $k^{th}$ *gene of the* $i^{th}$ *quantum moth*
$|sin(\theta_{i,k})|^2$*the probability of selecting* $k^{th}$ *gene of the* $i^{th}$ *quantum moth*

Simultaneously, the flame population containing *n* and the number *d* of Q-bits. It is represented as $QF = \{Qf_1, Qf_2, \dots, Qf_d\}$ where $Qf_j$ ($j$ = 1, 2, ... $n$) is the $j^{th}$ flame. Each $Qf_j$ quantum flame represents as follows (Eq.3):

$$Qf_i = \begin{vmatrix} \cos(\omega_{j,1}) & \cos(\omega_{j,2}) & \cos(\omega_{j,3}) & \dots\dots & \cos(\omega_{j,d}) \\ \sin(\omega_{j,1}) & \sin(\omega_{j,2}) & \sin(\omega_{j,3}) & \dots\dots & \sin(\omega_{j,d}) \end{vmatrix}$$

Where, $\omega_{j,k}$ ($k$ = 1, ....., $d$) satisfy the normalization condition
$|cos(\omega_{i,k})|^2 + |sin(\omega_{i,k})|^2 = 1$ and $\omega_{i,k} \epsilon [0, \frac{\pi}{2}]$

### 2.2. The Hybrid QMFOA Approach

The basic structure of the QMFOA has presented in this paper is described by Algorithm 1.

## 3. RESULTS AND DISCUSSIONS

Accuracy is one of the evaluation criteria of the classification model. The accuracy of the classification is the overall correctness of the classifier and is defined as the sum of the true correct cancer classifications divided by the total number of classifications. The accuracy of the classification is calculated according to Eq. 4.

$$Classification\ Accuracy = \frac{CC}{N} \times 100$$

Where, $N$ is the total number of the instances in the initial Microarray dataset and $CC$ refers to correct classified instances.

### 3.1. Dataset

Table 1 presents detailed characteristics of these gene expression datasets in terms of the number of classes, the number of genes, sample size, reference, and a brief description.

### 3.2. Parameter Settings

The parameters used in QMFOA are displayed in Table 2. The QMFOA used the mMRM as pre-filters to select the 100 top-ranked genes from all. In this study, to perform our experiments, the number of runs is 10 times on each dataset.

### 3.3. Experimental Results and Analysis

The QMFOA is evaluated on two kinds of benchmark Microarray cancer data, which are binary class and multi class datasets, in order to evaluate the performance and prove the effectiveness of the QMFOA to the gene selection problem. To achieve this, we made a couple of comparisons with some recently published algorithms.

---

**Algorithm 1:** QMFOA pseudo-code.

**Input:**
*Data*: Dataset; ▷ Data set
*dsubset*: set of genes; ▷ The set of genes of genes by pre-filter mRMR (d: genes).
*Pop_Size*: integer; ▷ population size.
*QM* [*Pop_Size*], *QF* [*Pop_Size*] of Quantum individual;
▷ *QM*: Quantum moth and *QF*: Quantum flame.
*BM* [*Pop_Size*], *BF* [*Pop_Size*] of Binary individual;
▷ *BM*: Binary moth and *BF*: Binary flame.
*FM* [*Pop_Size*], *FF* [*Pop_Size*] of Float;
▷Fitness of moth population *FM* and flame population *FF*.

**Output:**
*Subsetbest* : set of genes; ▷ Best subset of genes

1: Normalization_Min-Max (Data); ▷Normalization of the dataset
2: *dsubset* ←mRMR (*Data*); ▷ the subset of genes that obtained by pre-filter of the mRMR
3: **for** ($k \leftarrow 1$ **to** *Pop_Size*) **do**
4: *Initialization* ($QM^0_{i,k}, |d_{subset}|$ ▷ Initialize a population $QM^0$ of $|d_{subset}|$quantum moths
5: **end for**
6: *Iteration* ← 0
7: **repeat**
8: *N br_Flames* is calculated using by [16])
9: **for** ($i \leftarrow 1$ **to** *Pop_Size*) **do**
10: $BM_i^t$ ←Transformation ($QM_i^t, |d_{subset}|$); ▷Make a $BM^t$ of $|d_{subset}|$ from $QM^t$ by Transformation function that defined in Algorithm 2
11: *FM* [i] ←Evaluate $BM_i^t$ to fitness function;
12: **end for**
13: **if** *Iteration* ← 0 **then**
14: *QF; BF; FF* ← *sort* ($FM^t$, $QM^t$, $BM^t$); ▷Sort $QM^t$, $BM^t$ in ascending order by the fitness function / ($t$ = 0)
15: **else**

16: $QF;\ BF;\ FF \leftarrow sort\ (FM^t;\ [QF;QM^t];\ [BF;BM^t])$; ▷ Sort $[QF;QM^t]$, $[BF;BM^t]$ in ascending order by the fitness function i.e., among moth population ($t-1$, $t$)
17: **end if**
18: **for** ($i \leftarrow 1$ **to** *Pop_Size*) **do**
19: **for** ($j \leftarrow 1$ **to** *N br_Flames*) **do**
20: $Dist_{i;j} \leftarrow Distance(\ BM_i^t\ ,\ BF_j^t$ Calculate the distance between the $i^{th}$ moth ($BM_i^t$) , the $j^{th}$ flame ($BF_j^t$ )
21: **end for**
22: $Qm_i^{t+1} \leftarrow Update\_Q\ (Qm_i^t, Dist_{i;j}\ , |d_{subset}|)$ ▷Update the quantum moth by *Update_Q* function that defined in Algorithm 3
23: **end for**
24: $Subset_{best} \leftarrow QF_0, BF_0, FF$ [0]
25: $Iteration \leftarrow Iteration + 1$
26: **until** (*Iteration* >*Max_iteration*)
27: **return** $Subset_{best}$

### *3.3.1. Results obtained by QMFOA on binary class datasets*

In this work, we compare QMFOA with well-known gene selection algorithms published in the literature that have applied to binary datasets, which present in Table 3, like PCC-BPSO and PCC-GA [12], MOBBA_LS [6], GBC [3], ICA-ABC [4], MIM-AGA [15].
Table 3 illustrates the comparison of the experimental results between the QMFOA and other gene selection algorithms that applied to binary datasets, in terms of the best, worst, average and standard deviation (S.D.) of the number of genes selected and the classification accuracy. Cells with unknown values, to our knowledge, are represented with the '-' character.

**Algorithm 2:** Transformation pseudo-code.
**Input:**
$Qm_i$ : quantum individual; ▷$Qm_i$: quantum moth.
$d$: integer; ▷ where, d: the maximum number of genes in a moth

**Output:**
$Bm_i$ : binary individual; ▷ $Bm_i$: binary moth.

1: **function** *Transformation* ($Qm_i$, $d$)
2: **for** ($j \leftarrow 1$ **to** $d$) **do**
3: $threshold \leftarrow random\ (0, 1)$ ▷ Generate a random real value between 0 and 1
4: **if** ($threshold > |Cos\ (\theta_{i,j})|^2$ **then**
5: $Bm_{i;j} \leftarrow 1$ ;
6: **else**
7: $Bm_{i;j} \leftarrow 0$;
8: **end if**
9: **end for**
10: **return** $Bm_i$
11: **end function**

**Algorithm 3:** Updating-Quantum pseudo-code.
**Input:**
$Qm_i^t$ : quantum moth at generation $t$;
$Dist_{i;j}$ : real; ▷ $Dist_{i,j}$ : Distance between $i^{th}$ moth and $j^{th}$ flame at generation $t$.
$d$: integer; ▷ where $d$: the maximum number of genes in the moth $|d_{subset}|$.

**Output:**
$Qm_i^{t+1}$ : quantum moth of the next generation (t+1);
$BM_i^{t+1}$ : binary moth of the next generation (t+1);

1: **function** UPDATE_Q ($Qm_i^t$, $Dist_{i;j}$ , $d$)
2: **for** ($k \leftarrow 1$ **to** $d$) **do**
3: **if** $fitness\ (Bm_{i;k}) < fitness(Bf_{j;k})$ **then**
4: $\alpha \leftarrow rand\ (0, 2 * k * \pi)$
5: $\theta_{i,k}^{t+1} \leftarrow Dist_{i;j} * \alpha$
6: **else**
7: $\theta_{i,k}^{t+1} \leftarrow \theta_{i,k}^t$
8: **end if**
9: **end for**
10: $Bm_i^{t+1} \leftarrow Transformation\ (Qm_i^{t+1}, d)$ ▷Apply *Transformation* on a new quantum moth $Qm_i^{t+1}$
11: **return** $Bm_i^{t+1}$, $Qm_i^{t+1}$
12: **end function**

As can be seen in Table 3, for Leukemia1, four algorithms (PCC-BPSO, PCC-GA, GBC and our algorithm) can obtain 100% classification accuracy. For MOBBA_LS, ICA-ABC, and MIM-AGA methods selected 3, 5 and 7 genes and achieved 97.1%, 96.43%, and 97.68% classification accuracy, respectively. In contrast, our algorithm selects 32 genes and achieves 100% classification accuracy. In the best-obtained results for Leukemia1, the QMFOA obtained a slightly larger amount of genes than PCC-BPSO and GBC. For Prostate_Tumor, our method has achieved the highest accuracy (100%), but the MOBBA_LS method selected 6 genes and achieved 94.10% classification accuracy, in terms of accuracy, it is very far from QMFOA, which is better than all methods.

In addition, for CNS, the QMFOA method has achieved the highest accuracy that is better than all methods by 100%, at all classification accuracy. For Colon, the ICA-ABC method selected 12 genes and achieved 90.22% average accuracy.

**Table 1: Description for the test gene expression datasets.**

| Dataset Name | Samples | Features | Classes | Notes | Source |
|---|---|---|---|---|---|
| CNS | 60 | 7129 | 2 (Binary class) | 'MS': 39, 'TF': 21 | [27] |
| Colon | 62 | 2000 | 2 (Binary class) | 'Tumor': 40, 'Normal': 22 | [2] |
| Leukemia1 | 72 | 7129 | 2 (Binary class) | 'ALL': 47, 'AML': 25 | [11] |
| Breast | 97 | 24481 | 2 (Binary class) | 'non-relapse': 51, 'relapse': 46) | [27] |
| Ovarian | 253 | 15154 | 2 (Binary class) | 'Cancer': 162, 'Normal': 91 | [22] |
| Prostate_Tumor | 102 | 10509 | 2 (Binary class) | 'Normal': 52, 'Tumor': 50 | [24] |

**Table 2: QMFOA parameters for gene subset selection and classification.**

| **parameters** | **Setting value** |
|---|---|
| Population size | 50 |
| Normalization interval | [-1,1] |
| $\omega_1$ | 0.70 |
| $\omega_2$ | 0.30 |
| Random angle (Archimedes spiral) | $[0, 6\pi]$ / $k = 3$ |
| Number of generation (iteration) | 30 |
| Top-ranked genes by mMRM | 100 |

In contrast, the QMFOA is better than all methods; it selects 30.67 genes and achieves 100% classification accuracy. For Breast, the MIM-AGA method selected 216 genes and achieved 95.21% average accuracy. But, our algorithm selects 27.73 average genes and achieves 81.44% classification accuracy. On the other hand, the QMFOA selected 0.12% of the genes in terms of the number of genes selected by the MIM-AGA.
Finlay, For the Ovarian dataset, the PCC-BPSO, PCC-GA, and QMFOA can provide 100% in terms of best accuracy with the number of genes selected being 17, 22, and 17, respectively. The QMFOA can provide more than 99% for average accuracy and less than 20 selected genes.
Based on the above analysis, in this comparison, we can conclude that QMFOA has given better results than other algorithms in terms of the classification accuracy and the number of genes selected.

**Table 3: Comparison of experimental results obtained by MIM-mMFA with other methods for binary class datasets**.

| Algorithms | | Dataset | Leukemia1 | Prostate Tumor | CNS | Colon | Breast | Ovarian |
|---|---|---|---|---|---|---|---|---|
| QMFOA | Accuracy | Best | 100,00 | 100,00 | 100,00 | 100,00 | 81,44 | 100,00 |
| | | Worst | 100,00 | 98,02 | 100,00 | 100,00 | 74,23 | 98,42 |
| | | Avg. | 100,00 | 99,87 | 100,00 | 100,00 | 77,53 | 99,37 |
| | | S.D. | 0,00 | 0,51 | 0,00 | 0,00 | 2,07 | 0,44 |
| | # Genes | Best | 32,00 | 26,00 | 28,00 | 27,00 | 22,00 | 17,00 |
| | | Worst | 41,00 | 39,00 | 40,00 | 34,00 | 33,00 | 24,00 |
| | | Avg. | 36,47 | 32,60 | 31,27 | 30,67 | 27,73 | 20,60 |
| | | S.D. | 3,00 | 4,14 | 3,63 | 2,09 | 2,96 | 2,16 |
| PCC-BPSO | Accuracy | Best | 100,00 | 97,06 | 98,33 | 91,94 | 90,72 | 100,00 |
| | # Genes | Best | 18,00 | 33,00 | 39,00 | 25,00 | 41,00 | 17,00 |
| PCC-GA | Accuracy | Best | 100,00 | 96,08 | 98,33 | 91,94 | 88,66 | 100,00 |
| | # Genes | Best | 35,00 | 26,00 | 48,00 | 29,00 | 38,00 | 22,00 |
| MOBBA_LS | Accuracy | Best | 97,10 | 94,10 | - | - | - | - |
| | # Genes | Best | 3,00 | 6,00 | - | - | - | - |
| GBC | Accuracy | Best | 100,00 | - | - | 98,38 | - | - |
| | | worst | 93,05 | - | - | 91,93 | - | - |
| | | Avg | 96,43 | - | - | 94,62 | - | - |
| | # Genes | Best | 5,00 | - | - | 20,00 | - | - |
| ICA-ABC | Accuracy | Best | 98,21 | 97,88 | - | 97,34 | - | - |
| | | worst | 55,76 | 77,81 | - | 82,34 | - | - |
| | | Avg | 83,22 | 82,34 | - | 90,22 | - | - |
| | # Genes | Best | 12,00 | 20,00 | - | 12,00 | - | - |
| MIM-AGA | Accuracy | Best | 97,68 | 97,69 | - | 89,09 | 95,21 | - |
| | # Genes | Best | 7,00 | 93,00 | - | 19,00 | 216,00 | - |

## 4. CONCLUSIONS

In this work, we have presented a new hybrid technique between quantum computing and the Moth flam optimization called quantum Moth flam optimization algorithm (QMFOA) for gene selection and classification of high dimensional datasets. Therefore, the goal of this work is to provide a new bio-inspired algorithm to solve gene selection problems.

The QMFOA consists of two stages. In the first stage, we used the mRMR as a pre-filter method to rank the gene scores and select the 100 top genes as inputs of the second stage. The overall objective of this work is to select a smaller number of genes and obtain a classification accuracy similar to or better than that obtained by using all genes.

The experimental results of QMFOA on six binary class datasets have shown that our algorithm can find useful informative genes than all other compared algorithms in terms of classification accuracy, i.e. is better than all other compared algorithms, and also able to deliver competitive results in terms of the number of genes.

## REFERENCES

[1] Abderrahim, A., Talbi, E.G., Khaled, M., 2012. Hybridization of genetic and quantum algorithm for gene selection and classification of microarray data. International Journal of Foundations of Computer Science 23, 431–444.

[2] Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., Levine, A.J., 1999. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proceedings of the National Academy of Sciences 96, 6745–6750.

[3] Alshamlan, H.M., Badr, G.H., Alohali, Y.A., 2015. Genetic bee colony (gbc) algorithm: A new gene selection method for microarray cancer classification. Computational biology and chemistry 56, 49–60.

[4] Aziz, R., Verma, C., Jha, M., Srivastava, N., 2017. Artificial neural network classification of microarray data using new hybrid gene selection method. International Journal of Data Mining and Bioinformatics 17, 42–65.

[5] Cover, T.M., Van Campenhout, J.M., 1977. On the possible orderings in the measurement selection problem. IEEE transactions on systems, man, and cybernetics 7, 657–661.

[6] Dashtban, M., Balafar, M., Suravajhala, P., 2018. Gene selection for tumor classification using a novel bio-inspired multi-objective approach. Genomics 110, 10–17.

[7] Davies, S., Russell, S., 1994. Np-completeness of searches for smallest possible feature sets, in: AAAI Symposium on Intelligent Relevance, AAAI Press. pp. 37–39.

[8] Deng, X., Li, Y., Weng, J., Zhang, J., 2019. Feature selection for text classification: A review. Multimedia Tools and Applications 78, 3797–3816.

[9] Du, D., Li, K., Li, X., Fei, M., 2014. A novel forward gene selection algorithm for microarray data. Neurocomputing 133, 446–458.

[10] Duval, B., Hao, J.K., Hernandez Hernandez, J.C., 2009. A memetic algorithm for gene selection and molecular classification of cancer, in: Proceedings of the 11th Annual conference on Genetic and evolutionary computation, ACM. pp. 201–208.

[11] Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., et al., 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. science 286, 531–537.

[12] Hameed, S.S., Muhammad, F.F., Hassan, R., Saeed, F., 2018. Gene selection and classification in microarray datasets using a hybrid approach of pcc-bpso/ga with multi classifiers. JCS 14, 868–880.

[13] Han, K.H., Kim, J.H., 2002. Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. IEEE transactions on evolutionary computation 6, 580–593.

[14] Jain, A., Nandakumar, K., Ross, A., 2005. Score normalization in multimodal biometric systems. Pattern recognition 38, 2270–2285.

[15] Lu, H., Chen, J., Yan, K., Jin, Q., Xue, Y., Gao, Z., 2017. A hybrid feature selection algorithm for gene expression data classification.

[16] Mirjalili, S., 2015. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowledge-Based Systems 89, 228–249.

[17] Mudaliar, P.U., Patil, T.A., Thete, S.S., Moholkar, K.P., 2015. A fast clustering based feature subset selection algorithm for high dimensional data. International journal of emerging trend in engineering and basic science 2, 494–499.

[18] Mundra, P.A., Rajapakse, J.C., 2010. Gene and sample selection for cancer classification with support vectors based t-statistic. Neurocomputing 73, 2353–2362.

[19] Narayanan, A., 1999. Quantum computing for beginners, in: Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406), IEEE. pp. 2231–2238.

[20] Nielsen, M.A., Chuang, I., 2002. Quantum computation and quantum information.

[21] Peng, H., Long, F., Ding, C., 2005. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and minredundancy. IEEE Transactions on Pattern Analysis & Machine Intelligence , 1226–1238.

[22] Petricoin III, E.F., Ardekani, A.M., Hitt, B.A., Levine, P.J., Fusaro, V.A., Steinberg, S.M., Mills, G.B., Simone, C., Fishman, D.A., Kohn, E.C., et al., 2002. Use of proteomic patterns in serum to identify ovarian cancer. The lancet 359, 572–577

[23] Sardana, M., Agrawal, R., Kaur, B., 2016. A hybrid of clustering and quantum genetic algorithm for relevant genes selection for cancer microarray data. International Journal of Knowledge-based and Intelligent Engineering Systems 20, 161–173.

[24] Singh, D., Febbo, P.G., Ross, K., Jackson, D.G., Manola, J., Ladd, C., Tamayo, P., Renshaw, A.A., D'Amico, A.V., Richie, J.P., et al., 2002. Gene expression correlates of clinical prostate cancer behavior. Cancer cell 1, 203–209.

[25] Wang, H., Niu, B., 2017. A novel bacterial algorithm with randomness control for feature selection in classification. Neurocomputing 228, 176–186.

[26] Xi, M., Sun, J., Liu, L., Fan, F., Wu, X., 2016. Cancer feature selection and classification using a binary quantum-behaved particle swarm optimization and support vector machine. Computational and mathematical Methods in Medicine 2016.

[27] Zhu, Z., Ong, Y.S., Dash, M., 2007. Markov blanket-embedded genetic algorithm for gene selection. Pattern Recognition 40, 3236–3248.