International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 4 (2019), pp. 457-464
© International Research Publication House. http://www.irphouse.com

A hybrid genetic-fuzzy-rough-mutual information (GFRMI) based method for cancer classification

Thilagavathy Chenniappan[1], **Rajesh Reghunadhan**[2]

[1]*CMS College of Science and Commerce, Coimbatore, Tamilnadu, India.*

[2]*Department of Computer Science, Central University of Kerala,*
Periya – 671316, Kasaragod, Kerala, India.

Abstract

Classification of cancer microarray dataset is one of the greatest challenge, due the presence of more number genes (features). Recently, Fuzzy sets/logic, rough sets and mutual information has independently made a major change in the dimensionality reduction. But still improvements are needed in terms of feature (dimensionality) reduction leading to a higher classification rate. The objective of this paper is to minimize the selected number of features and to maintain a higher classification rate, due to the life-critical nature of cancer microarray dataset. This paper presents a hybrid genetic-fuzzy-rough-mutual information (GFRMI) method for the effectual classification of cancer microarray data. The proposed GFRMI based method helps to make use of the best base methods for the gene selection for effectual classification of microarray cancer classification. The results on benchmarking microarray cancer datasets show that the proposed method provides better results with less number of genes.

Keywords: Microarray, cancer classification, genetic-fuzzy-rough-mutual information method, dimensionality reduction, gene selection.

I. INTRODUCTION

Precise identification or prediction of cancer-type from several hundred types of cancer is most important for proper treatment and therapy. Due to the limitations of biopsy methods (difficulty in knowing cells “growth-rate” [1], level of “penetration” [2], depth of “metastatic cascade” [3], and higher chances of development of “resistance towards agents” [4]) and due to the limitations of molecular methods utilizing RNA/DNA/Protein (difficulty in knowing biotic generation/progression of cancer), microarrays [5] developed by Patrick O. Brown et. al. in 1990’s based on the principle of “base-paring/hybridization” are considered to be one of the most relevant method which provides additional information/patterns concurrently from several thousands of genes for diagnosis/classification of cancer including its subtypes. A multi-category classification method by Statnikov et. al helped the world understand the importance of machine learning for microarray gene expression cancer diagnosis [6].

But, there is a high need of identifying the motif genes which are responsible for the disease which will further lead to an efficient forecasting/prediction method. One such method is the dimensionality reduction mechanism by making use of techniques like principle component analysis (PCA) [7], linear discriminant analysis (LDA) [8], etc. But, PCA/LDA still lack in exactly identifying the genes which are responsible for disease.

Several methods have been developed during the last few years for the selection of features/genes from the dense microarray data; and these selected genes can be utilized for the effectual classification of the cancer, quantitative measurements in “wet-lab” and for the easiest diagnosis from fluids/serum. Feature selection in the gene expression data using improved binary particle swarm optimization (PSO) is one of the important work in microarray data [9]. Fuzzy rule based binary PSO is another similar work for feature selection [10]. Lin et. al have proposed a method for selecting feature subsets based on support vector machine and recursive feature elimination (SVM-RFE) for the classification [11].

The other important works in this regard include, but not limited to, the hierarchical gene selection based on genetic-fuzzy-system [12] and feature selection based on fuzzy-rough-uncertainty metric [13]. A very dense understanding on feature selection methods can be had from the survey by Chandrashekar et. al [14]. Even though several methods were proposed for gene selection, still better one are needed for gene selection and also for improving the classification accuracy.

The theory of entropy and mutual information by Shannon (1940’s) had laid a foundation in almost all disciplines for dealing with uncertainty and information content. The theory of fuzzy sets/logic by Zadeh (in 1965) have added additional feather for handling imprecise knowledge and for representing uncertainty/vagueness/ambiguity [15]. The introduction of rough sets in 1981 by Pawlak was another landmark and turning point in handling uncertainty in decision systems [16]. Many researchers started using these for information representation, reduction and prediction. The most important application among those are the feature reduction/selection using entropy, mutual information, rough sets, fuzzy sets and also hybrid integration of these techniques.

The importance of fuzzy logic and rough sets in gene selection was proved by Hu et. al in his works, namely, “information preserving hybrid data reduction based on fuzzy rough sets (IP-FRS)” [17], “fuzzy probabilistic approximation spaces (FPAS)” [18], “entropies of fuzzy indiscernibility relations (E-FIR)” [19] and “heterogeneous feature subset selection using neighborhood rough sets (NRS)” [20]. Each of the methods has its own advantages and disadvantages. The features selected by the respective methods were different and hence the results were not optimal.

Even a small improvement of the results based on some novel or hybrid method in life critical problems like cancer prediction is very important and most significant. Hence, the objective of this paper is to minimize the selected number of features and to maintain a higher classification rate, due to the life-critical nature of cancer microarray dataset. This paper makes use of the relative importance of each of the methods by Hu et. al. The paper proposes a hybrid fuzzy-rough-mutual information base method, in which, the feature subsets obtained from the methods based on mutual information, entropy, fuzzy information entropy, kernalized fuzzy rough sets, preference learning rough sets, preference learning fuzzy rough sets, neighbourhood rough sets (NRS), NRS with variable precision lower approximation (VPLA), and fuzzy NRS with VPLA are combined together to form a super-subset of selected features. Moreover, this paper also presents a genetic algorithm based gene selection using the super-subset of selected features for the effectual classification of cancer microarray data. The results on benchmarking microarray cancer datasets show that, the number of features selected is much lesser with better classification accuracy when compared to the other papers in the literature.

This paper is organized as follows. Section 2 presents materials and methods. Results and discussions are presented in Section 3. Section 4 concludes the paper.

II. MATERIALS AND METHODS

This section presents the proposed hybrid genetic-fuzzy-rough-mutual information method for the classification of microarray cancer dataset. It works in two different stages as shown in figure 1.

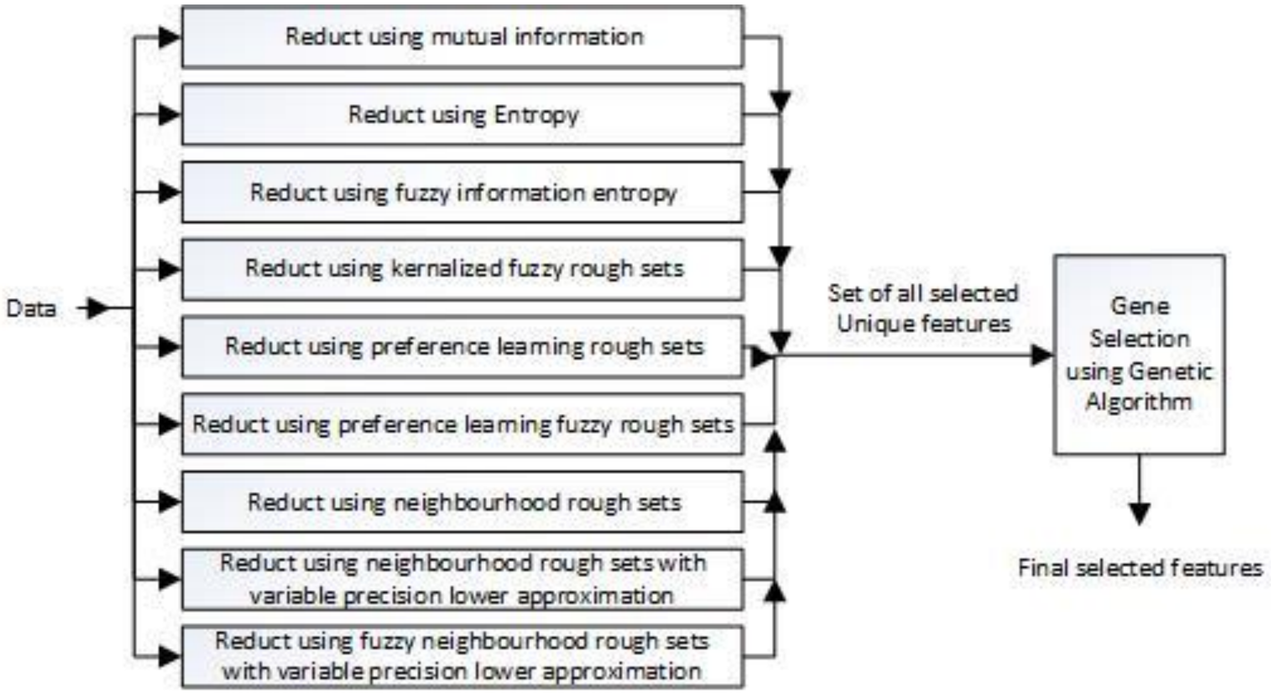
Data
Reduct using mutual information
Reduct using Entropy
Reduct using fuzzy information entropy
Reduct using kernalized fuzzy rough sets
Reduct using preference learning rough sets
Reduct using preference learning fuzzy rough sets
Reduct using neighbourhood rough sets
Reduct using neighbourhood rough sets with variable precision lower approximation
Reduct using fuzzy neighbourhood rough sets with variable precision lower approximation
Set of all selected Unique features
Gene Selection using Genetic Algorithm
Final selected features

Figure 1. Genetic algorithm based gene selection using hybrid fuzzy-rough-mutual information methods.

In the first stage nine different feature selection methods proposed by Hu et. al is used. The features selected by the respective methods were different and hence the results were not optimal. Since each of the methods proposed by Hu et. al has its own advantages and disadvantages, the fusion of all the feature subset provided by the nine different methods is utilized in this paper to get a super-feature subset which has the advantages of all the algorithms. In the second stage, a genetic algorithm is used to select the features form the unique features obtained from the previous step.

2.1. Stage 1 of hybrid genetic-fuzzy-rough-mutual information method

The sudo-codes of the most important methods utilized in this paper are provided in this section. Considering the importance of fuzzy logic, rough sets, entropy and mutual information, the methods are adopted from Qinghua Hu et.al. [17-20]. The feature selection based on information entropy is shown in *Algorithm 1*. Similarly, feature selection based on neighborhood rough sets, neighborhood rough sets with variable precision lower approximation, fuzzy information entropy, fuzzy neighborhood rough sets, fuzzy neighborhood rough sets with variable precision lower approximation are shown respectively in *Algorithm 2,3,4,5,6*. The following symbols and equations are used in the algorithms.

L → *the column vector of class labels of all records*

N → *the number of samples/records*

t_1 → *is a threshold*

F_i → *i^{th} feature (ith column/attribute) vector*

$R_{L} = abs(L_{matrix} - L_{matrix}^{T}) < t_1$ *where* $L_{matrix} = repmat(L, [1,N])$

$R_b = ones(N)$

$R_{LF} = (abs(L_{matrix} - L_{matrix}^{T}) < t_1).*(1- abs(L_{matrix} - L_{matrix}^{T})/ t_1)$

$E_L = entropy(R_L)$

Algorithm 1: Reduct using information entropy

```
for each column j of the attributes
    for each column i of the attributes
        $F=repmat(F_i, [1,N])$
        $R_i=abs(F - F^T) < t_1$
        $D(i)= E_L + E(min(R_i, R_b))-E(min(min(R_i, R_L), R_b));$
    end
    $[v(j),k]=max(D)$
    If $abs(v(j)-v(j-1))>t_2$
            $F=repmat(F_k, [1,N]);$
        $R_{nb} = abs(F - F^T) < t_1$
        $R_b=min(R_b, R_{nb});$
        save $k^{th}$ feature as one of the best feature
    else  break
end
```

Algorithm 2: Reduct using neighborhood rough sets

```
for each column j of the attributes
    for each column i of the attributes
        $F=repmat(F_i, [1,N])$
        $R_i=abs(F - F^T) < t_1$
        $r1=min(R_i, R_b);$
        $nn=r1'.*repmat(F_i, [1 row]);$
        for class_i=1:classnum
                C(class_i,:)=sum(nn==class_i);
        end
        [value,real_class]=max(C);
        $D(i)= sum(real\_class'== F_i)/row;$
    end
    $[v(j),k]=max(D)$
    If $abs(v(j)-v(j-1))>t_2$
        $F=repmat(F_k, [1,N]);$
        $R_{nb} = abs(F - F^T) < t_1$
        $R_b=min(R_b, R_{nb});$
        save $k^{th}$ feature as one of the best feature
    else        break
end
```

Algorithm 3: Reduct using Neighborhood rough sets with variable precision lower approximation

```
for each column j of the attributes
    for each column i of the attributes
        $F=repmat(F_i, [1,N])$
        $R_i=abs(F - F^T) < t_1$
        $r1=min(R_i, R_b);$
        $mr=min(r1, R_L);$
        incluse=sum(mr').\/sum(r1');
        DI=(incluse>=inclusion);
        MRR=(sum(mr').\/length(mr'~=0));
        D(i)= (DI.*MRR)/row;
    end
    $[v(j),k]=max(D)$
    If $abs(v(j)-v(j-1))>t_2$
        $F=repmat(F_k, [1,N]);$
        $R_{nb} = abs(F - F^T) < t_1$
        $R_b=min(R_b, R_{nb});$
        save $k^{th}$ feature as one of the best feature
    else      break
end
```

Algorithm 4: Reduct using fuzzy information entropy

```
for each column j of the attributes
    for each column i of the attributes
        F=repmat($F_i$, [1,N])
        $R_i=(abs(F - F^T) < t_1).*(1- abs(F - F^T)/ t_1)$
        $D(i)= E_L + E(min(R_i, R_b))-E(min(min(R_i, R_{LF}), R_b))$;
    end
    [v(j),k]=max(D)
    If abs(v(j)-v(j-1))>$t_2$
            F=repmat($F_k$, [1,N]);
        $R_{nb} = (abs(F – F^T) < t_1).* (1-abs(F – F^T) / t_1)$
        $R_b=min(R_b, R_{nb})$;
        save $k^{th}$ feature as one of the best feature
    else        break
end
```

Algorithm 5: Reduct using fuzzy neighborhood rough sets

```
for each column j of the attributes
    for each column i of the attributes
        F=repmat($F_i$, [1,N])
        $R_i=(abs(F - F^T) < t_1).*(1- abs(F - F^T)/ t_1)$
        r1=min($R_i$, $R_b$);
        nn=r1'.*repmat($F_i$, [1 row]);
        for class_i=1:classnum
            C(class_i,:)=sum(nn==class_i);
        end
        [value,real_class]=max(C);
        D(i)= sum(real_class'== $F_i$)/row;
    end
    [v(j),k]=max(D)
    If abs(v(j)-v(j-1))>$t_2$
        F=repmat($F_k$, [1,N]);
        $R_{nb} = (abs(F – F^T) < t_1).* (1-abs(F – F^T) / t_1)$
        $R_b=min(R_b, R_{nb})$;
            save $k^{th}$ feature as one of the best feature
    else      break
end
```

Algorithm 6: Reduct using fuzzy Neighborhood rough sets with variable precision lower approximation

```
for each column j of the attributes
    for each column i of the attributes
        F=repmat($F_i$, [1,N])
        $R_i=(abs(F - F^T) < t_1).*(1- abs(F - F^T)/ t_1)$
        r1=min($R_i$, $R_b$);
        mr=min(r1, $R_L$);
        incluse=sum(mr')./sum(r1');
        DI=(incluse>=inclusion);
        MRR=(sum(mr')./length(mr'~=0));
        D(i)= (DI.*MRR)/row;
    end
    [v(j),k]=max(D)
    If abs(v(j)-v(j-1))>$t_2$
        F=repmat($F_k$, [1,N]);
        $R_{nb} = (abs(F – F^T) < t_1).* (1-abs(F – F^T) / t_1)$
        $R_b=min(R_b, R_{nb})$;
            save $k^{th}$ feature as one of the best feature
    else      break
end
```

Other feature selection based on mutual information, feature selection based on kernelized fuzzy rough sets, feature selection based on preference learning based on rough sets and feature selection based on preference learning fuzzy rough sets are also used [17][18][19][20]. The algorithms are run on the dataset by having various combinations of values for the parameters in the algorithm and the features obtained and combined together to form the super-subset of features. The values of the parameters used in the algorithm are shown in table 1.

Table 1. Parameters and their values used in the algorithms

Parameters	Values
Neighborhood radius	0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5
Options	Fuzzy/Crisp
Inclusion	0.8, 0.85, 0.90, 0.95, 1
Evaluating measures for kernelized fuzzy rough sets	‘GD_S’ - dependency function based on S-T model ‘GD_theta’ - dependency function based on theta-eta model ‘GW_S’ - classification certainty function based on S-T model ‘GW_theta’ - classification certainty function based on theta-sigma model
delta (the kernel parameter)	0.01, 0.1, 1, 2
K (the number of the nearest samples to compute the evaluating measure)	1, 2, 3, 4, 5, 10
Evaluating measure for Preference learning based on fuzzy rough sets	'FUC' - upwards consistency, 'FLC' - downwards consistency, 'FGC' - global consistency
Evaluating measure for Preference learning based on rough sets	'UC' - upwards consistency, 'LC' - downwards consistency, 'GC' - global consistency

Parameters for heterogeneous greedy feature selection based on neighborhood rough sets.	Delta = [0.1,0.2]
	efc_ctrl=0.01,0.1,0.2,1,2,10

2.2. Stage 2 – Genetic algorithm based feature selection

Genetic algorithm (GA) is a simple, powerful, derivative free optimization technique derived from the Darwinian’s theory of evolution subject to selection, crossover/recombination and mutation [21]. The major advantage is that GA evolves and progress with multiple solution points in the search-space/domain. The operators can be applied independently to a solution or a pair of solutions. Thus GA’s are capable to run on multiple/distributed/clustered systems.

A set of features that need to be optimized can be considered as genes. These genes together will form a chromosome. Each chromosome is a solution point in the search domain. A set of chromosome forms a population.

Initially the GA starts with the random initialization of the genes in the chromosome. Once initialized, we can find out the fitness/goodness of each of the solution/chromosome in the population. The goodness/fitness can be found out by using a fitness function (fitness function will be different for each problem). For example, in the case of 3D modeling of an aircraft, the fitness function may be inversely proportional to the air friction.

Then GA runs through a loop consisting of operators like selection, crossover and mutation. Each loop completes one generation. The aim of selection mechanism is to make sure that good chromosomes with more fitness are selected with higher chance/probability. The simplest selection function is the Roulette wheel selection, in which each chromosome is assigned a proportionate area in the wheel based on the percentage of fitness. The roulette wheel is pointed by the pointer will be chosen. According to the probability theory, there are higher chances of selection for those chromosomes which are having higher span of area. The selection mechanism ensures the increase in the average fitness of the population from generation to generation.

The crossover function will take to chromosomes and then exchange a set of randomly selected genes. The crossover is carried out with a crossover-probability. There are various types of crossovers, namely, single point crossover, double point crossover, multipoint crossover, shuffle crossover, arithmetic crossover, donation based crossover, sharing based crossover, etc.

The mutation operator will take a chromosome and will random modify the values inside the gene with a mutation-probability. The mutation probability will be very less to ensure that there is no wide destruction in the chromosomes. GA has wide applications across the disciplines including, but not limited to, fuzzy system optimization, travel salesman problem, gene selection etc.

100 binary chromosomes are use in proposed algorithm in which ‘1’ indicate the selection of the concerned feature. The crossover rate and mutation rate are selected as 0.8 and 0.2 respectively. A 50% reinsertion rate is used as part of elitism. The classification results of k-nearest neighborhood method are used as fitness values of the chromosomes.

2.3. Datasets

The dataset used in the study are from the benchmarking and widely used GEMS (a system for automated cancer diagnosis and biomarker discovery from microarray gene expression data) [6]. As an initial study to test the performance of the hybrid method, small round blood cell tumors (SRBCT) dataset is used with the following childhood types, namely, “Ewings sarcoma”, “neuroblastoma”, “Burkitt’s lymphoma” and “rhabdomyosarcoma”.

Then three cancer datasets are considered, namely, the Leukemia dataset (dataset 1 and dataset 2 with 5327 and 11225 features respectively and consisting of 72 samples) and Lung cancer dataset (with 12600 features and 203 samples). Table 2 shows the details of the datasets used in this paper.

Table 2. Datasets

Dataset	Sample count	Feature length	Class types
SRBCT	83	2308	4
Leukemia1	72	5327	3
Leukemia2	72	11225	3
Lung Cancer	203	12600	5

3. RESULTS AND DISCUSSIONS

The results of proposed hybrid method on the datasets, namely, SRBCT, Leukemia1, Leukemia2, and Lung cancer dataset are shown respectively in Table 3, 4, 5, 6.

Table 3. Comparison of results on SRBCT dataset

Method	Selected features	Classification accuracy (%)
KNN by Statnikov et. al [6]	2308	86.90
Kernelized fuzzy rough set (KFRS) + Transductive (Semisupervised) SVM (TSVM) by Chakraborty et. al [22]	-	98.06
Fuzzy rule based particle swarm optimization (FRBPSO) by Agarwal et. al [10]	213	98.19
Multiclass support vector machine (MC-SVM) by Statnikov et. al [6]	2308	100
improved binary particle swarm optimization (IBPSO)+KNN by Chuang et. al 9	431	100
Proposed	137	100

Table 4. Comparison of results on Leukemia1 dataset

Method	Selected features	Classification accuracy (%)
KNN by Statnikov et. al [6]	5327	83.57
Multiclass support vector machine (MC-SVM) by Statnikov et. al [6]	5327	97.50
Fuzzy rule based particle swarm optimization (FRBPSO) by Agarwal et. al [10]	825	98.19
Improved binary particle swarm optimization (IBPSO)+KNN by Chuang et. al [9]	1034	100
Proposed method	101	100

Table 5. Comparison of results on Leukemia2 dataset.

Method	Selected features	Classification accuracy (%)
KNN by Statnikov et. al [6]	11225	87.14
M-SVM-RFE-OA [11]	73.99	94.69±2.03
multiclass support vector machine (MC-SVM) by Statnikov et. al [6]	11225	97.32
fuzzy rule based particle swarm optimization (FRBPSO) by Agarwal et. al [10]	1028	97.50
improved binary particle swarm optimization (IBPSO)+KNN by Chuang et. al [9]	1292	100
Proposed method	206	100

Table 6. Comparison of results on Lung Cancer dataset.

Method	Selected features	Classification accuracy (%)
KNN by Statnikov et. al [6]	12600	89.64
Multiclass support vector machine (MC-SVM) by Statnikov et. al [6]	12600	96.55
Improved binary particle swarm optimization (IBPSO)+KNN by Chuang et. al [9]	1897	96.55
Proposed method	161	97.04

The results on all the four dataset show that the proposed method provides better classification accuracy with minimum number of features/genes when compared to other works in the literature. The proposed method has wide applicability and still there are rooms for further improvement in minimizing the number of features and at same time keeping the accuracy at the highest.

IV. CONCLUSION

In this paper a hybrid genetic-fuzzy-rough-mutual information method is proposed for the selection of minimal genes which are best enough to classify the caner effectively. The results on benchmarking cancer dataset show that the proposed method provides good classification accuracy with less number of features/genes.

Acknowledgements

First author would like to thank Bharathiar University and CMS College for their support in doing the research work. The second author would like to acknowledge the research support provided by Central University of Kerala

REFERENCES

[1] K.M. Kerr, D. Lamb, Actual growth rate and tumour cell proliferation in human pulmonary neoplasms, *Br J Cancer.*, 50(3), 1984, 343–349.

[2] S.I. Jeffrey, K. William, B. Robert, Tissue requirements in lung cancer diagnosis for tumor heterogeneity, mutational analysis and targeted therapies: initial experience with intra-operative Frozen Section Evaluation (FROSE) in bronchoscopic biopsies, *J. Thorac. Dis*., 8(6), 2016, S488–S493.

[3] A.M. Tracey, Y. Lin, J.S. Andrew, L. Jane, G.J. Wen, Cancer Invasion and Metastasis: Molecular and Cellular Perspective. In: Metastatic Cancer: Clinical and Biological Perspectives, J. Rahul (Eds.). *Landes Bioscience*, 2013

[4] H. Kim, K.J. Chae, S.H. Yoon, M. Kim, B. Keam, T.M. Kim, D.W. Kim, J.M. Goo, C.M. Park, Repeat biopsy of patients with acquired resistance to EGFR TKIs: implications of biopsy-related factors on T790M mutation detection. *Eur Radiol*., 28(2), 2018, 861-868.

[5] P.O. Brown, D. Botstein, Exploring the new world of the genome with DNA microarrays. Nat. Genet., 21, 1999, 33-37.

[6] A. Statnikov, C.F. Aliferis, I. Tsamardinos, D. Hardin, S. Levy, A Comprehensive evaluation of multicategory

classification methods for microarray gene expression cancer diagnosis, *Bioinformatics*, 21(5), 2005, 631-643.

[7] E. Lotfi, A. Keshavarz, Gene expression microarray classification using PCA-BEL, *Computers in Biology and Medicine*, 54, 2014, 180-187.

[8] E.B. Huerta, B. Duval and J.K. Hao, A hybrid LDA and genetic algorithm for gene selection and classification of microarray data. *Neurocomputing*, 73(13-15), 2010, 2375-2383.

[9] L.Y. Chuang, H.W. Chang, C.J. Tu, C.H. Yang, Improved binary PSO for feature selection using gene expression data. *Computational Biology and Chemistry*, 32, 2008, 29-38.

[10] S. Agarwal, R. Rajesh, P. Ranjan, FRBPSO: A fuzzy rule based binary PSO for feature selection. *Proceedings of the National Academy of Sciences, India (Section A: Physical Sciences),* 87(2), 2017, 221-233.

[11] X. Lin, C. Li, Y. Zhang, B. Su, M. Fan, H. Wei, Selecting feature subsets based on SVM-RFE and the overlapping ratio with applications in bioinformatics, *Molecules*, 23(1), 2017, 52.

[12] T. Nguyen, A. Khosravi, D. Creighton, S. Nahavandi, Hierarchical gene selection and genetic fuzzy system for cancer microarray data classification. *PLOS ONE*, 10(3), 2015, e0120364.

[13] J. Xu, Y. Wang, K. Xu, T. Zhang, Feature genes selection using fuzzy rough uncertainty metric for tumor diagnosis, *Computational and Mathematical Methods in Medicine*, Aricle ID 6705648, 2019.

[14] G. Chandrashekar, F. Sahin, A survey on feature selection methods. Computers and electrical engineering, 40, 2014, 16-28.

[15] L.A. Zadeh, Fuzzy logic = computing with words, *IEEE transaction on fuzzy systems*, 4(2), 1996, 103-111.

[16] Z. Pawlak, Z., Rough sets, *International Journal of Parallel Programming*, 11(5), 1982, 341-356

[17] Q.H. Hu, D.R. Yu, Z. Xie, Information preserving hybrid data reduction based on fuzzy-rough techniques. Pattern recognition letters, 27(5), 2006, 414-423.

[18] Q.H. Hu, D.R. Yu, Z. Xie, J. Liu, Fuzzy probabilistic approximation spaces and their information measures, *IEEE Transactions on fuzzy systems*, 14(2), 2006, 191-201.

[19] Q.H. Hu, D.R. Yu, Entropies of fuzzy indiscernibility relation and its operations, *International Journal of uncertainty fuzziness and knowledge-based systems*, 12(5), 2004, 575-589.

[20] Q.H. Hu, D.R. Yu, J. Liu, C. Wu, Neighborhood rough set based heterogeneous feature subset selection, *Information Sciences*, 178, 2008, 3577-3594.

[21] D.E. Goldberg, *Genetic Algorithms*, 1st Edn., Pearson Education, 2009.

[22] D. Chakraborty, U. Maulik, Identifying Cancer Biomarkers from microarray data using feature selection and semi-supervised learning, *IEEE Journal of Translational Engineering in Health and Medicine*, 2, 2014.