



Pipelining the ranking techniques for microarray data classification: A case study



Rasmita Dash^{a,*}, Dr. Bijan Bihari Misra^b

^a Department of Computer Sc. & Information Technology, Institute of Technical Education and Research, Siksha 'O' Anusandhan, University, Khandagiri Square, Bhubaneswar, 751030 Odisha, India

^b Department of Computer Sc. & Engineering, Silicon Institute of Technology, Bhubaneswar, 751024 Odisha, India

ARTICLE INFO

Article history:

Received 7 August 2015

Received in revised form 23 March 2016

Accepted 3 July 2016

Available online 20 July 2016

Keywords:

Microarray data

Feature selection

Feature ranking technique

Classification

Statistical test

ABSTRACT

Identification of relevant genes from microarray data is an apparent need in many applications. For such identification different ranking techniques with different evaluation criterion are used, which usually assign different ranks to the same gene. As a result, different techniques identify different gene subsets, which may not be the set of significant genes. To overcome such problems, in this study pipelining the ranking techniques is suggested. In each stage of pipeline, few of the lower ranked features are eliminated and at the end a relatively good subset of feature is preserved. However, the order in which the ranking techniques are used in the pipeline is important to ensure that the significant genes are preserved in the final subset. For this experimental study, twenty four unique pipeline models are generated out of four gene ranking strategies. These pipelines are tested with seven different microarray databases to find the suitable pipeline for such task. Further the gene subset obtained is tested with four classifiers and four performance metrics are evaluated. No single pipeline dominates other pipelines in performance; therefore a grading system is applied to the results of these pipelines to find out a consistent model. The finding of grading system that a pipeline model is significant is also established by Nemenyi post-hoc hypothetical test. Performance of this pipeline model is compared with four ranking techniques, though its performance is not superior always but majority of time it yields better results and can be suggested as a consistent model. However it requires more computational time in comparison to single ranking techniques.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

In recent years, analysis of microarray gene expression data comes up as a challenging issue in bioinformatics research. In this context, classification of microarray samples represents a well-studied problem in statistics and machine learning, where a large number of successful methods have been suggested [1]. However, it has also been shown that commonly used baseline classifiers pose intrinsic draw-backs in achieving accurate and reproducible results. In order to obtain more robust microarray data analysis, sophisticated methods should be applied for classification and prediction of microarray data [2].

Microarray data set or gene expression data sets are organized as matrix form and is experimented on different samples. The column represents different genes in gene expression data and row repre-

sents sample measured at different time point. There are number of gene expression data analysis techniques and classification is one out of them. In classification a classifier will assign a new sample to one of the existing class.

For example in a two class problem, a sample is assigned to disease affected class (positive class) or normal class (negative class). Again in microarray data the number of genes is too high (in the range of 1000–10000) and number of sample is comparatively low (in the range of 100). This, however, poses a great challenge to traditional classification algorithms. So this high dimension increases the search space and makes the classification task more complex. So gene selection or feature selection is prerequisite for classification. For efficient high dimensional data classification Dash and Liu [3] carried out a comprehensive overview of feature selection techniques. There are different feature selection methods and they are categorised based on some criterion. For example feature selection methods may be categorised into two ways called as filter and wrapper method [4,5].

One of the principal selection mechanisms in different feature selection method is feature ranking. That is the features are ranked

* Corresponding author.

E-mail addresses: rasmitadash@soauniversity.ac.in, rasmita02@yahoo.co.in (R. Dash), misrabijan@gmail.com (Dr.B.B. Misra).

based on some merit score computed based on ranking criterion and then few top ranked features are selected. It is a filter approach and has advantage over other feature selection scheme like it is simple to use, computationally and statistically scalable to large datasets and it has wide range of application.

Feature ranking method selects the significant features based on some selection criterion such as distance [6,7], information theory [8] or on some function of classifier's output. For example RELIEFF, which was proposed by Kira and Rendell [6] is one of the most successful distance based measure and adopt Euclidean distance to assign a relevance weight to each feature. But whatever may be the criterion, all conventional rank selection algorithms are based on single approach of evaluating the features. The difference in approach generates different ranks of features by different ranking methods. Therefore appropriate selection of features depends on the approach of evaluation of the ranking technique, which may favour the search space of the problem or may not. However, given a problem at hand, one does not possess the a priori knowledge about which criterion works the best for it. Therefore selection of appropriate ranking technique for the problem is an important as well as a difficult task.

This paper proposes pipelining of ranking techniques for efficient feature selection and classification of microarray databases. Four feature ranking methods are considered here for pipelining, such as information gain (IG), signal to noise ratio (SNR), Pearson correlation coefficient (PCC) and t-statistic. Based on the combination of 4 ranking methods, 24 pipelines are built up. Performance of the pipelines are evaluated using 4 classifiers, such as Multiple Linear Regression (MLR), Artificial Neural Network (ANN), Naïve Bayes network (NB) and k-Nearest Neighbour (k-NN). Further to find out which pipeline shows better performance, a grading method is used. In the first stage grading, the pipelines are graded with respect to different classifiers but result shows that the grading value varies from dataset to dataset. So it is difficult to conclude which pipeline performs the best. However, some of the pipelines perform better for most of the datasets, though not the best always. Therefore, grading in multiple stages is applied and the findings are validated using nonparametric statistical test.

Rest part of the paper is arranged as follows. In Section 2 recent literature studies on feature ranking method applied on microarray data is presented. In Section 3 the proposed model for feature selection and classification is discussed. The complete experimental work is described in Section 4. It includes the dataset used for the experimental analysis, data normalization process, dataset training, testing and validation process, feature ranking techniques used, performance matrices and result analysis. Finally the paper is concluded with Section 5.

2. Background study

The curse of dimensionality in microarray dataset makes it essential to have dimensionality reduction of the datasets before proceeding for classification. In this section different feature selection techniques are discussed with special reference to different ranking methods with their advantages and limitations.

2.1. Different feature selection techniques

Feature selection (or variable elimination) techniques are broadly categorised into two groups filter method and wrapper method [9]. The filter method considers the intrinsic properties of the data and generates a relevant set of features [10]. In this technique the genes are either ranked (depending on some scoring value) [11] or evaluated with respect to the cost function to identify the most informative genes [12]. The rank based filter method

Table 1
List of ranking techniques applied to different microarray database.

Author	Feature Ranking Method	Reference
Hall et al.	Information Gain	[16]
Peyman et al.	t-test, ANOVA	[17]
M. A. Hall et al.	Correlation based feature selection (CBF)	[18]
Thomas et al., Tsai et al.	t-test	[19,20]
Thomas et al., Antoniadis et al.	Wilcoxon score test	[21,22]
Hwang et al.	Wilks's Lambda score	[23]
Wang et al. Golub et al.	Signal to noise ratio	[24,25]
Cho et al. Ho et al.	Euclidian distance	[26,27]
Xing et al.	Information Gain	[28]

generates the relevant feature based on the intrinsic property of the dataset. In this technique features are evaluated on the basis of a scoring function where the top ranked features are selected and low ranked features are removed [11]. So to select the relevant feature a threshold value of the scoring function is chosen. The features with greater than or equal to the score value are selected and rest are removed. After that the subset features are presented for classification. The filter techniques are mostly effective as they are computationally fast and are easily scale to very high dimension data [13].

2.2. Feature selection using ranking methods in microarray databases

Recently, the ranked based feature selection has obtained more attention for solving the feature selection problem in many areas like sequence analysis, mass spectra analysis, single nucleotide polymorphisms (SNPs) analysis, text and literature mining, microarray data analysis and many more [12].

In many applications single ranking techniques applied to different microarray database listed in Table 1. Filter method rank each feature according to some univariate metric and only the highest ranking features are used while the remaining low ranking features are eliminated. This method also relies on general characteristics of the training data to select some features without involving any learning algorithm. Therefore, the results of filter model will not affect any classification algorithm. Moreover, filter methods also provide very easy way to calculate and can simply scale to large-scale microarray datasets since it has a short running time. In addition, filter methods also offer less computational time to generate results which is an extra point to be preferred by domain experts. However, gene ranking based on these techniques has some drawbacks. The major one is the genes selected are most probably redundant [14]. This is due to a rank method ranks the genes using single ranking criteria (or score function). But it is difficult to say which criteria suit a particular dataset to rank all the genes. Due to which some important genes may be rejected and some less important genes may be selected. To overcome this problem many embedded and ensemble gene selection techniques are applied to microarray databases [15].

Peng et al. [28] applied a hybrid approach using Fisher's ratio, a simple method easy to understand and implement, to filter out most of the irrelevant genes, then a wrapper method is employed to reduce the redundancy. The performance of FR-Wrapper approach is evaluated over four widely used microarray datasets (Leukemia, Lung cancer, Breast cancer and Colon cancer). Analysis of experimental results reveals that the hybrid approach can achieve the goal of maximum relevance with minimum redundancy.

A filter based on the t-statistic is used by Mundra et al. [29] in which t-statistic is divided into two parts, corresponding to *relevant* and *irrelevant* data points. The *relevant* data points are selected

using support vectors and then used to compute t -statistic for feature selection. By simultaneously selecting data points and genes, significantly better classification results are achieved using four t -statistic based methods are applied t -test, SVM based t -statistics, SVM with recursive feature elimination (RFE), and SVM based t -statistic with RFE on Leukemia database and shown 96.88%, 98.12%, 97.88%, and 98.41% accuracy.

An initial gene selection method used Lee et al. [30] implementing between groups to within groups (BW ratio) in order to improve the efficiency of microarray data analysis. Though it reduces the dataset, the screening of genes is done on per gene basis. It is not clear which and how many genes should be selected. So 500 genes are selected initially and then proposed algorithm genetic algorithm with dynamic parameter settings (GADP) is used to reduce the dataset with final selected genes. This technique is applied on six microarray database Colon cancer, small and round blue cell tumour (SRBCT), Breast cancer, ALL/AML, DLBCL and GCM database.

Abeel et al. [31] worked on the analysis of the robustness of biomarker selection techniques. A large-scale analysis of the recently introduced concept of ensemble feature selection is introduced, where multiple feature selections techniques are combined in order to increase the robustness of the final set of selected features. The selection methods that are embedded in the estimation of support vector machines (SVMs). SVMs are powerful classification models show state-of-the-art performance on several diagnosis and prognosis tasks on biological data. The proposed methodology is evaluated on four microarray datasets showing increases of up to almost 30% in robustness of the selected biomarkers, along with an improvement of approximately 15% in classification performance. The stability improvement with ensemble methods is particularly noticeable for small signature sizes (a few genes), which is most relevant for the design of a diagnosis or prognosis model from a gene signature.

Yang et al. [32] made an exhaustive analysis on the robustness issue existing in feature selection for high-dimensional and proposed a multi-criterion fusion-based recursive feature elimination (MCF-RFE) algorithm. It is developed with the goal of improving both classification performance and stability of feature selection results. Experimental studies on five gene-expression data sets show that the MCF-RFE algorithm outperforms the commonly used benchmark feature selection algorithm SVM-RFE.

3. Proposed work

In general different ranking techniques assign different rank to a gene depending on the evaluation scheme of the ranking technique. Therefore the top ranked genes in one method may not be the top ranked genes in another method. However, each ranking technique has got its significance for some problem space but not for others. Therefore one ranking technique may perform the best for one type of problem where others may not.

In this study, instead of selecting a small subset of genes by one ranking technique, the merits of all the ranking techniques are considered. When the set of genes pass through a pipeline of ranking techniques, the first ranking technique ranks the genes and rejects few of the lowest ranked genes and passes the rest genes to the next ranking technique in the pipeline. Subsequent ranking techniques in the pipeline also apply the same principle one after the other. The genes selected using a pipeline of feature ranking method are analysed using different classification techniques.

The basic design of the model considered for this case study is presented in Fig. 1. In the first step the microarray database is normalized and then it is presented to different pipelines. After selection of genes in the pipeline, the database is prepared for the cross validation. Then for each fold of the cross validation the

Table 2
Description of microarray databases.

Database	Genes	Samples		
		In Class1	In Class2	Total
Leukemia [34]	7129	27	11	38
Lungcancer1 [35]	7129	86	10	96
Lungcancer2 [36]	12533	15	134	149
CNS [37]	7129	39	21	60
Prostate [38]	12600	52	50	102
Colon cancer [39]	2000	40	22	62
DLBCL Stanford [40]	4026	24	23	47

respective training set and test set are presented to the classifiers for training and testing. The result obtained from each fold of the data is then exposed for performance analysis of the pipelines. In Fig. 1, P1 to P24 represents the twenty four pipelines considered for the case study. In each pipeline there are four ranking techniques used which are represented by R1 to R4 in the order as presented in Fig. 1. Further the specific ranking techniques taken for this case study are represented as R1: SNR, R2: PCC, R3: Information Gain, R4: T statistic.

From experimental study it was observed that any sequence of ranking techniques in the pipeline does not yield same level of performance. The performance of a pipeline model changes as the classifier changes or the datasets changes. A single pipeline does not dominate all other pipelines when different classifiers on different dataset are tested. Therefore it is essential to perform statistical analysis to find a suitable pipeline which may not yield best performance always, but its overall performance is better than all other pipelines as well as the performance of any single ranking technique.

Earlier Demsar et al. [33] has ranked/graded classifiers for multiple datasets and has statistically established that one of the classifiers can be considered better than others, though it does not yield best performance always. For this experiment one stage of ranking/grading is good enough to find the required output. But for this experiment one stage of grading is not good enough to find a single pipeline superior over others when it is applied on multiple classifiers for multiple datasets with multiple pipelines for multiple performance indexes. Therefore a multistage grading is inevitable.

4. Empirical procedure

4.1. Microarray databases used for the case study

The effectiveness and efficiency of this method is verified using seven different gene expression databases. The descriptions of the databases are presented in Table 2. Microarray data are high dimensional data where the number of samples or observation are too less in comparison to number of genes or attributes. Samples are in the range of hundreds and attributes are in thousands. The microarray databases considered here for this experimentation belongs to two class problem.

4.2. Data normalization

In general the range of the feature values of the raw data varies widely in different databases. The objective function of many training algorithm do not work properly when range of feature values differ. If the objective function of the training algorithm uses distance between two records, then the distance is dominated by the feature with broad range of values, which misleads the objective function. Similarly the gradient descent based back propagation algorithm converges faster when the feature values are from the same range. To eliminate the influence of one factor over another

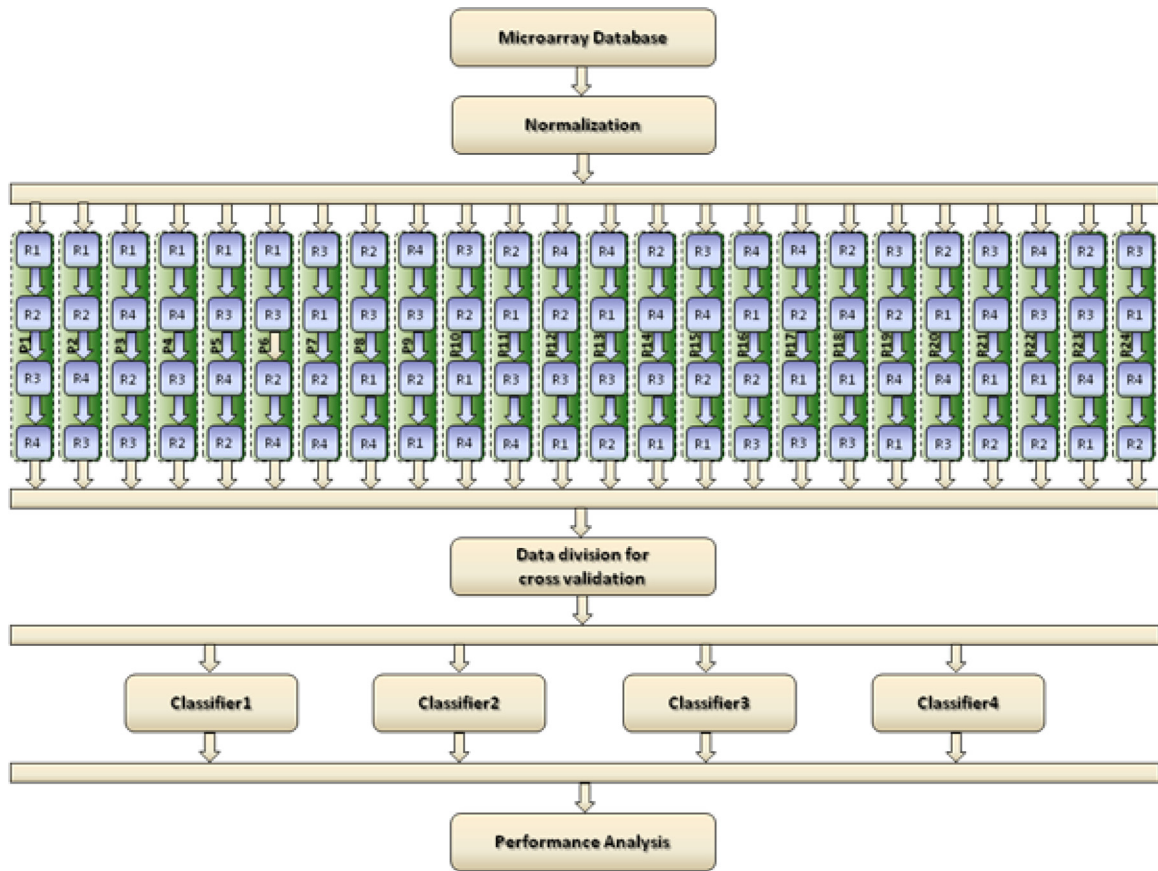


Fig. 1. Layout of the pipelined rank based microarray data classification.

and also for faster convergence, the features are scaled to a specified range. The process of scaling the feature values to a specified range is referred as data normalization. In microarray data the feature value ranges from a very small value to a very large value. Therefore data normalization is inevitable for microarray databases before use of any training algorithm. Here the min-max normalization method is applied to scale the data within the range of 0–1 as in Eq. (1).

$$\text{Normalized value} = \frac{\text{original value} - \min(\text{original value in the series})}{\max(\text{original value in the series}) - \min(\text{original value in the series})} \quad (1)$$

4.3. Generation of data for tenfold cross validation

In this experimental work to avoid issues of over fitting, tenfold cross validation technique has been applied. In this technique the original samples are divided into ten random partitions of equal size and each partition is formed in such a way that each partition contains roughly equal number of patterns per class. Out of these ten partitions, nine partitions are used for training and one partition is used for testing in each fold of the cross validation. In this scheme two different training sets differ by 10% of records.

4.4. Feature ranking methods used for pipeline

In this case study, four feature ranking techniques are considered for constructing the pipelines. Brief background of the feature ranking techniques is presented here.

4.4.1. Ranking methods

The proposed method uses filter technique for feature ranking. Few top ranked features are selected based on the statistical approach which represents enough information and functionality to enhance the performance of the classification. The statistical approach reduces the number of genes in the microarray data to avoid redundancy and misleading features. In this proposed work integration or pipelining of feature ranking method is considered. The gene ranking methods employed for pipelining are presented here.

4.4.1.1. *Signal to noise ratio (SNR)*. In this method genes are first ranked according to their expression levels using SNR test [41,42]. The SNR is defined in Eq. (2).

$$\text{SNR}(i) = (\mu_{i1} - \mu_{i2}) / (\sigma_{i1} - \sigma_{i2}) \quad (2)$$

where μ_{i1} is the mean value of samples belongs to of class 1 in i^{th} gene, μ_{i2} is the mean value of samples belongs to of class 2 in i^{th} gene, σ_{i1} is the standard deviations of the samples of class 1 in i^{th} gene, and σ_{i2} is the standard deviations of the samples of class 2 in i^{th} gene.

4.4.1.2. *Information gain (IG)*. Information gain is the expected reduction in entropy caused by partitioning the examples according to a given attribute.

$$S = \text{Info}(X) - \text{Info}_X(X) \quad (3)$$

$$\text{Info}(X) = - \sum_{i=1}^k P(c_i, X) * \log(P(c_i, X)) \quad (4)$$

Table 3
Number of genes selected from pipeline output.

Dataset	Number of gene selected
Leukemia	5
DLBCL Stanford	5
Prostate	4
Lungcancer1	5
Lungcancer2	4
CNS	9
Colon Cancer	3

$$Info(X) = - \sum_{i=1}^v \frac{|V_i|}{|X|} * Info(V_i) \quad (5)$$

Where k is the number of classes, v represents number of individual values of a gene x , v_i is the set of instances whose values in gene x is equal to x_i [43].

4.4.1.3. Pearson correlation coefficient (PCC). The correlation coefficient (CC) test is non-parametric test; it measures the correlation of genes with other samples. This is achieved by assuming an ideal gene named Y which is interpreted as the most informative gene that predicts the two classes of all the samples perfectly. Then the CC is calculated according to Eq. (6) where n_s is the total number of samples, Y is the ideal gene and X is the gene expression value. The highest the CC value the more informative the gene as it will be more correlated to ideal gene [41,44].

$$CC(i) = \frac{n_s \sum XY - \sum X \sum Y}{\sqrt{n_s \sum X^2 - (\sum X)^2 (n_s \sum Y^2 - (\sum Y)^2)}} \quad (6)$$

4.4.1.4. T-test. A t -test is a standard statistical approach for ranking differentially expressed genes from microarray data. General formulation of the t -test statistic for gene g is

$$t_g = \frac{m_1 - m_2}{\sqrt{\frac{s_1}{n_1} + \frac{s_2}{n_2}}} \quad (7)$$

Where m_1 and s_1 are the mean and the variance of the first group of patients, of size n_1 , and m_2 and s_2 are the mean and the variance of the second group of patients, of size n_2 [43].

4.5. Gene selection by the pipeline

In Section 3, Fig. 1 presents the proposed model, where twenty four pipelines each with four ranking techniques are used for selection of gene subset from the gene expression databases. The different ranking techniques employed here are signal to noise ratio, Pearson correlation coefficient, information gain, and t -statistic. Let the ranking techniques in the pipeline are R1, R2, R3, R4 in order from beginning of the pipeline to the end of the pipeline. The genes are first exposed to R1 for ranking, once the genes are ranked, 20% of the lowest ranked genes are removed and rest of the genes are presented to R2. The same process is repeated with R2, R3 and R4 i.e. at each stage 20% of the lowest ranked genes are removed from the dataset. However, after removal of the genes by R4, large numbers of genes remain in the non-rejected gene set. Using the rank of R4, few best ranked genes are considered for the purpose of classification. When a uniform number of genes are taken from all the data set, they do not show good result. Therefore different numbers of genes are selected from different data sets, which are obtained by performing number of experiments with different datasets. Table 3 shows the number of genes selected from different datasets after the pipeline operation.

Table 4
Parameter and parameter value in ANN.

Parameter in ANN	Parameter Value
Number of neuron in the Input layer	Number of genes in the reduced dataset
Number of neuron in the Hidden layer	4
Number of neuron in the Output layer	1
Learning rate	0.6
Momentum Coefficient	0.7

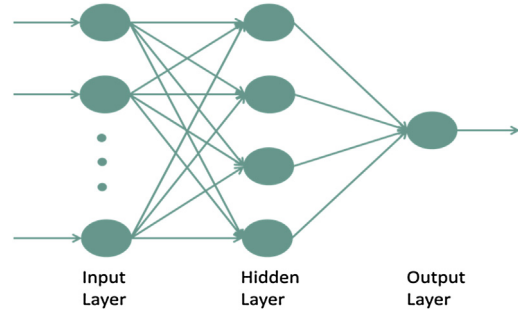


Fig. 2. ANN Neural Network Model.

4.6. Classifiers used for the case study

To generalize the objective of the selection through pipeline of ranking techniques, commonly and frequently used classifiers are considered here for the case study. Further a set of classifiers are considered here to avoid the specialized decision biased towards a classifier than that of a generalized decision which is more or less suitable for most of the classifiers. Therefore, in this experimental work, four classifiers such as multiple linear regression (MLR), artificial neural network (ANN), naïve Bayesian network (NB) and k -Nearest Neighbor (k -NN) are used to evaluate the performance of different pipeline models. In k -nearest neighbour since only one parameter i.e. the k value (number of neighbour) needs to be determined is considered as the most simple classifier. It is the most popular nonparametric method introduced by Fix and Hodges in 1951 [45]. Here the test sample is assigned to a class mostly closest to its k nearest neighbour depending in the nearest distance. Here Euclidian distance is employed to evaluate the distance between two microarray patterns.

For this case study, single layered feed forward neural network is used [46], the input layer fed by the number of genes selected for each dataset. Linear activation function is used in the input layer; sigmoidal activation function is used in the hidden and output layers [47–49]. So number of neuron in the input layer varies from database to database. The number of genes selected for the input layer being small, only four neurons are considered for the hidden layer and one neuron for the output layer as shown in Fig. 2. The parameter setting for the ANN considered here is presented in Table 4.

The motivation of using naïve Bayes classifier is its simplicity, robustness and also it works efficiently for noisy and irrelevant attributes. It assumes each feature contribute independently of the probability to classify a sample belong to a particular class. Despite its assumption and its naïve designing, it works efficiently in many real world situations [50].

MLR is one of the classical statistical models that represent a quantitative relationship between a dependent and independent variable of the dataset. Multiple linear regression is an extension of linear regression involving more than one predictable variable. MLR attempts to model the relationship between two or more explanatory variables and a response variable by fitting linear equation to

Table 5
Performance and ranking of Leukemia microarray database.

Pipeline	Accuracy				Sensitivity				Specificity				Jaccard				Overall Average Grade
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	
P1	0.94 (8.5)	1.00 (7)	1.00 (4.5)	0.94 (3)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.94 (3.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	5.78
P2	0.95 (3)	1.00 (7)	0.98 (15)	0.92 (8)	0.90 (15)	1.00 (7)	0.95 (16)	0.70 (9)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.92 (8)	0.90 (15)	1.00 (7)	0.95 (16)	0.70 (8)	10.59
P3	0.95 (3)	1.00 (7)	0.98 (15)	0.92 (8)	0.90 (15)	1.00 (7)	0.95 (16)	0.70 (9)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.92 (8)	0.90 (15)	1.00 (7)	0.95 (16)	0.70 (8)	10.59
P4	0.94 (5)	0.99 (15.5)	0.98 (15)	0.88 (14)	0.90 (15)	0.97 (15.5)	0.95 (16)	0.65 (13)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.89 (13)	0.90 (15)	0.97 (15.5)	0.95 (16)	0.63 (13)	13.56
P5	0.92 (24)	0.99 (15.5)	0.98 (15)	0.88 (14)	0.90 (15)	0.97 (15.5)	0.95 (16)	0.65 (13)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.89 (13)	0.90 (15)	0.97 (15.5)	0.95 (16)	0.63 (13)	14.75
P6	0.94 (8.5)	1.00 (7)	1.00 (4.5)	0.94 (3)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.94 (3.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	5.22
P7	0.94 (8.5)	1.00 (7)	1.00 (4.5)	0.94 (3)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.94 (3.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	5.78
P8	0.94 (8.5)	1.00 (7)	1.00 (4.5)	0.94 (3)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.94 (3.5)	0.99 (15)	1.00 (7)	1.00 (5)	0.80 (3.5)	6.44
P9	0.92 (17.5)	0.67 (22)	0.98 (15)	0.85 (20)	0.90 (15)	0.15 (22)	0.95 (16)	0.50 (20)	0.98 (14)	0.74 (22)	1.00 (11.5)	0.85 (19.5)	0.90 (15)	0.15 (22)	0.95 (16)	0.50 (20)	17.97
P10	0.98 (1)	1.00 (7)	1.00 (4.5)	0.88 (12)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.60 (15)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.88 (15)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.60 (15)	8.03
P11	0.94 (8.5)	1.00 (7)	1.00 (4.5)	0.94 (3)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.94 (3.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.80 (3.5)	5.78
P12	0.92 (17.5)	0.67 (22)	0.94 (24)	0.85 (20)	0.80 (23.5)	0.15 (22)	1.00 (5)	0.50 (20)	0.94 (24)	0.74 (22)	1.00 (11.5)	0.85 (19.5)	0.80 (23.5)	0.15 (22)	1.00 (5)	0.50 (20)	15.91
P13	0.94 (8.5)	0.99 (15.5)	0.98 (15)	0.94 (3)	0.90 (15)	0.97 (15.5)	0.95 (16)	0.80 (3.5)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.94 (3.5)	0.9 (15)	0.97 (15.5)	0.95 (16)	0.80 (3.5)	9.38
P14	0.92 (17.5)	0.67 (22)	0.95 (23)	0.85 (20)	0.80 (23.5)	0.15 (22)	0.90 (23)	0.50 (20)	0.95 (23)	0.74 (22)	0.98 (23.5)	0.85 (19.5)	0.80 (23.5)	0.15 (22)	0.90 (23)	0.50 (20)	21.72
P15	0.92 (17.5)	0.67 (22)	0.98 (15)	0.85 (20)	0.90 (15)	0.15 (22)	0.95 (16)	0.50 (20)	0.98 (14)	0.74 (22)	1.00 (11.5)	0.85 (19.5)	0.90 (15)	0.15 (22)	0.95 (16)	0.50 (20)	17.97
P16	0.92 (17.5)	1.00 (7)	1.00 (4.5)	0.90 (10.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.70 (9)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.90 (10.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.65 (10.5)	8.03
P17	0.92 (17.5)	1.00 (7)	0.98 (15)	0.85 (20)	0.90 (15)	1.00 (7)	0.95 (16)	0.50 (20)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.85 (19.5)	0.90 (15)	1.00 (7)	0.95 (16)	0.50 (20)	14.41
P18	0.92 (17.5)	1.00 (7)	0.98 (15)	0.85 (20)	0.90 (15)	1.00 (7)	0.95 (16)	0.50 (20)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.85 (19.5)	0.90 (15)	1.00 (7)	0.95 (16)	0.50 (20)	14.41
P19	0.92 (17.5)	1.00 (7)	1.00 (4.5)	0.90 (10.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.70 (9)	1.00 (4.5)	1.00 (10)	1.00 (11.5)	0.90 (10.5)	1.00 (4.5)	1.00 (7)	1.00 (5)	0.65 (10.5)	8.03
P20	0.95 (3)	1.00 (7)	0.98 (15)	0.92 (8)	0.90 (15)	1.00 (7)	0.95 (16)	0.70 (9)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.92 (8)	0.90 (15)	1.00 (7)	0.95 (16)	0.70 (8)	10.59
P21	0.92 (17.5)	0.98 (18.5)	0.98 (15)	0.85 (20)	0.90 (15)	0.95 (18.5)	0.95 (16)	0.50 (20)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.85 (19.5)	0.90 (15)	0.95 (18.5)	0.95 (16)	0.50 (20)	14.06
P22	0.92 (17.5)	0.99 (15.5)	0.96 (22)	0.88 (14)	0.87 (22)	0.97 (15.5)	0.87 (24)	0.65 (13)	0.98 (14)	1.00 (10)	0.98 (23.5)	0.89 (13)	0.87 (22)	0.97 (15.5)	0.87 (24)	0.63 (13)	15.47
P23	0.92 (17.5)	0.67 (22)	0.98 (15)	0.85 (20)	0.90 (15)	0.15 (22)	0.95 (16)	0.50 (20)	0.98 (14)	0.74 (22)	1.00 (11.5)	0.85 (19.5)	0.90 (15)	0.15 (22)	0.95 (16)	0.50 (20)	17.97
P24	0.92 (17.5)	0.98 (18.5)	0.98 (15)	0.85 (20)	0.90 (15)	0.95 (18.5)	0.95 (16)	0.50 (20)	0.98 (14)	1.00 (10)	1.00 (11.5)	0.85 (19.5)	0.90 (15)	0.95 (18.5)	0.95 (16)	0.50 (20)	16.56

Table 6
Performance and ranking of Colon microarray databases.

Pipeline	Accuracy				Sensitivity				Specificity				Jaccard				Overall Average Grade
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	
P1	0.68 (17)	0.8 (4.5)	0.77 (1)	0.77 (13.5)	0.25 (17)	0.69 (4.5)	0.62 (4)	0.62 (12.5)	0.7 (17)	0.83 (4.5)	0.78 (1)	0.78 (12.5)	0.18 (17)	0.54 (4.5)	0.47 (12.5)	0.47 (13)	9.75
P2	0.73 (8.5)	0.74 (16.5)	0.64 (16)	0.81 (3.5)	0.51 (8.5)	0.6 (17)	0.55 (7)	0.65 (5)	0.78 (8.5)	0.77 (17)	0.64 (16)	0.8 (3)	0.44 (7.5)	0.5 (16.5)	0.38 (7)	0.5 (8)	10.34
P3	0.73 (8.5)	0.74 (16.5)	0.68 (10)	0.81 (3.5)	0.51 (8.5)	0.6 (17)	0.25 (16.5)	0.65 (5)	0.78 (8.5)	0.77 (17)	0.7 (4.5)	0.8 (3)	0.44 (7.5)	0.5 (16.5)	0.18 (16.5)	0.52 (4.5)	10.22
P4	0.68 (17)	0.8 (4.5)	0.73 (5)	0.77 (13.5)	0.25 (17)	0.69 (4.5)	0.68 (2)	0.62 (12.5)	0.7 (17)	0.83 (4.5)	0.69 (7)	0.78 (12.5)	0.18 (17)	0.54 (4.5)	0.47 (2.5)	0.47 (13)	9.63
P5	0.68 (17)	0.8 (4.5)	0.75 (2.5)	0.77 (13.5)	0.25 (17)	0.69 (4.5)	0.5 (10)	0.62 (12.5)	0.7 (17)	0.83 (4.5)	0.74 (3)	0.78 (12.5)	0.18 (17)	0.54 (4.5)	0.33 (10)	0.47 (13)	10.19
P6	0.68 (17)	0.8 (4.5)	0.64 (16)	0.77 (13.5)	0.25 (17)	0.69 (4.5)	0.55 (7)	0.62 (12.5)	0.7 (17)	0.83 (4.5)	0.64 (16)	0.78 (12.5)	0.18 (17)	0.54 (4.5)	0.38 (7)	0.47 (13)	11.47
P7	0.75 (6)	0.8 (4.5)	0.61 (22)	0.77 (13.5)	0.58 (4)	0.69 (4.5)	0.28 (13)	0.62 (12.5)	0.78 (5)	0.83 (4.5)	0.65 (14)	0.78 (12.5)	0.44 (5)	0.54 (4.5)	0.21 (13)	0.47 (13)	7.91
P8	0.75 (6)	0.8 (4.5)	0.58 (24)	0.77 (13.5)	0.58 (4)	0.69 (4.5)	0.27 (14)	0.62 (12.5)	0.78 (5)	0.83 (4.5)	0.59 (19)	0.78 (12.5)	0.44 (5)	0.54 (4.5)	0.17 (19)	0.47 (13)	10.34
P9	0.53 (20.5)	0.56 (22)	0.66 (12)	0.61 (22)	0.19 (21)	0 (22)	0 (23)	0.28 (22)	0.65 (20.5)	0.6 (22)	0.66 (11)	0.64 (22)	0.13 (21)	0 (22)	0 (232)	0.21 (22)	22.00
P10	0.75 (6)	0.8 (4.5)	0.62 (23)	0.77 (13.5)	0.58 (4)	0.69 (4.5)	0.06 (20)	0.62 (12.5)	0.78 (5)	0.83 (4.5)	0.57 (20)	0.78 (12.5)	0.44 (5)	0.54 (4.5)	0.05 (20)	0.47 (13)	10.10
P11	0.68 (17)	0.8 (4.5)	0.68 (9)	0.77 (13.5)	0.25 (17)	0.69 (4.5)	0.25 (16.5)	0.62 (12.5)	0.7 (17)	0.83 (4.5)	0.7 (4.5)	0.78 (12.5)	0.18 (17)	0.54 (4.5)	0.18 (16.5)	0.47 (13)	11.50
P12	0.5 (23)	0.56 (22)	0.66 (14)	0.61 (22)	0.19 (21)	0 (22)	0 (23)	0.28 (22)	0.62 (22)	0.6 (22)	0.66 (11)	0.64 (22)	0.13 (21)	0 (22)	0 (23)	0.21 (22)	20.88
P13	0.79 (2)	0.78 (12)	0.71 (8)	0.74 (19)	0.57 (6)	0.63 (4.5)	0.26 (15)	0.55 (19)	0.82 (1)	0.81 (12)	0.65 (13)	0.75 (19)	0.47 (3)	0.5 (16.5)	0.17 (18)	0.37 (19)	11.69
P14	0.53 (20.5)	0.56 (22)	0.66 (12)	0.61 (22)	0.19 (21)	0 (22)	0 (23)	0.28 (22)	0.65 (20.5)	0.6 (22)	0.66 (11)	0.64 (22)	0.13 (21)	0 (22)	0 (23)	0.21 (22)	20.50
P15	0.5 (23)	0.56 (22)	0.63 (20)	0.61 (22)	0.13 (23.5)	0 (22)	0.24 (18.5)	0.28 (22)	0.58 (23.5)	0.6 (22)	0.43 (22)	0.64 (22)	0.07 (23.5)	0 (22)	0.2 (14.5)	0.21 (22)	21.53
P16	0.68 (12.5)	0.74 (16.5)	0.63 (18.5)	0.81 (3.5)	0.43 (12)	0.4 (17)	0.4 (11.5)	0.65 (5)	0.76 (12)	0.77 (17)	0.37 (23.5)	0.8 (3)	0.39 (11)	0.5 (16.5)	0.28 (11.5)	0.52 (4.5)	11.50
P17	0.68 (12.5)	0.76 (13)	0.63 (18.5)	0.81 (3.5)	0.43 (12)	0.63 (12.5)	0.4 (11.5)	0.65 (5)	0.76 (12)	0.79 (13.5)	0.3723.5 (3)	0.8 (3)	0.39 (11)	0.52 (11)	0.28 (11.5)	0.52 (4.5)	9.69
P18	0.78 (4)	0.76 (13)	0.64 (16)	0.79 (7)	0.55 (7)	0.63 (12.5)	0.55 (7)	0.6 (17.5)	0.78 (2)	0.79 (13.5)	0.64 (16)	0.79 (7)	0.38 (14)	0.52 (11)	0.38 (7)	0.5 (8)	10.16
P19	0.68 (12.5)	0.74 (16.5)	0.66 (12)	0.81 (3.5)	0.43 (12)	0.6 (17)	0.24 (18.5)	0.67 (1.5)	0.76 (12)	0.77 (17)	0.46 (21)	0.79 (6)	0.39 (11)	0.5 (16.5)	0.2 (14.5)	0.54 (1)	12.03
P20	0.68 (12.5)	0.74 (16.5)	0.71 (7)	0.81 (3.5)	0.43 (12)	0.6 (17)	0.58 (5)	0.65 (5)	0.76 (12)	0.7 (17)	0.69 (9)	0.8 (3)	0.39 (11)	0.5 (16.5)	0.35 (9)	0.52 (4.5)	10.03
P21	0.75 (2)	0.79 (10)	0.73 (5)	0.77 (13.5)	0.6 (1.5)	0.68 (10)	0.68 (2)	0.6 (17.5)	0.78 (5)	0.81 (10)	0.69 (7)	0.77 (18)	0.48 (1.5)	0.52 (11)	0.47 (2.5)	0.42 (18)	8.41
P22	0.75 (2)	0.79 (10)	0.75 (2.5)	0.77 (13.5)	0.6 (1.5)	0.68 (10)	0.53 (9)	0.62 (12.5)	0.78 (5)	0.81 (10)	0.74 (2)	0.78 (12.5)	0.48 (1.5)	0.52 (11)	0.41 (5)	0.47 (13)	7.56
P23	0.5 (23)	0.56 (22)	0.62 (23)	0.61 (22)	0.13 (23.5)	0 (22)	0.05 (21)	0.28 (22)	0.58 (23.5)	0.6 (22)	0.62 (8)	0.64 (22)	0.07 (23.5)	0 (22)	0.02 (21)	0.21 (22)	21.41
P24	0.68 (10)	0.79 (10)	0.73 (5)	0.79 (8)	0.43 (12)	0.68 (10)	0.68 (2)	0.67 (1.5)	0.76 (12)	0.81 (10)	0.69 (7)	0.78 (12.5)	0.39 (11)	0.52 (11)	0.47 (2.5)	0.52 (4.5)	8.06

Table 7
Performance and ranking of Prostate microarray database.

Pipeline	Accuracy				Sensitivity				Specificity				Jaccard				Overall Average Grade
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	
P1	0.967 (14)	0.91 (17)	0.89 (16.5)	0.85 (17)	0.87 (17.5)	0.87 (18)	0.87 (17.5)	0.83 (18)	0.91 (17.5)	0.98 (17)	0.91 (17.5)	0.85 (18)	0.73 (17.5)	0.85 (18)	0.73 (17.5)	0.82 (5)	16.47
P2	0.980 (4.5)	0.97 (10)	1.00 (4)	0.90 (12.5)	1.00 (5.5)	0.93 (12)	1.00 (5.5)	0.90 (13)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.89 (13.5)	1.00 (4)	0.93 (11)	1.00 (4)	0.70 (16.5)	8.41
P3	0.980 (4.5)	0.97 (10)	0.98 (11.5)	0.90 (12.5)	1.00 (5.5)	0.93 (12)	1.00 (5.5)	0.90 (13)	0.98 (12)	1.00 (7.5)	0.98 (12)	0.89 (13.5)	0.95 (12)	0.93 (11)	0.95 (12)	0.70 (16.5)	9.75
P4	0.951 (16.5)	0.94 (14.5)	0.95 (13.5)	0.93 (4.5)	0.97 (13)	0.93 (12)	0.97 (13)	1.00 (5)	0.96 (14.5)	0.98 (17)	0.96 (14.5)	0.92 (4.5)	0.87 (15.5)	0.88 (15.5)	0.87 (15.5)	0.80 (9.5)	12.41
P5	0.951 (16.5)	0.94 (14.5)	0.95 (13.5)	0.93 (4.5)	0.97 (13)	0.93 (12)	0.97 (13)	1.00 (5)	0.96 (14.5)	0.98 (17)	0.96 (14.5)	0.92 (4.5)	0.87 (15.5)	0.88 (15.5)	0.87 (15.5)	0.80 (9.5)	12.41
P6	0.967 (14)	0.91 (17)	0.71 (24)	0.85 (17)	0.52 (23)	0.87 (18)	0.52 (23)	0.83 (18)	0.79 (22)	0.98 (17)	0.79 (22)	0.85 (18)	0.40 (23)	0.85 (18)	0.40 (23)	0.82 (5)	18.88
P7	0.971 (9.5)	1.00 (2.5)	0.98 (9)	0.95 (2)	0.97 (13)	1.00 (3)	0.97 (13)	1.00 (5)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.94 (2)	0.97 (9)	1.00 (3)	0.97 (9)	0.85 (2)	6.28
P8	0.971 (9.5)	1.00 (2.5)	0.98 (9)	0.95 (2)	0.97 (13)	1.00 (3)	0.97 (13)	1.00 (5)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.94 (2)	0.97 (9)	1.00 (3)	0.97 (9)	0.85 (2)	5.34
P9	0.764 (21)	0.74 (22)	0.75 (20)	0.76 (21)	0.43 (24)	0.66 (22.5)	0.43 (24)	0.63 (22.5)	0.79 (23)	0.80 (22.5)	0.79 (23)	0.81 (21.5)	0.38 (24)	0.54 (22.5)	0.38 (24)	0.51 (21.5)	22.44
P10	0.971 (9.5)	1.00 (2.5)	0.98 (9)	0.95 (2)	0.97 (13)	1.00 (3)	0.97 (13)	1.00 (5)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.94 (2)	0.97 (9)	1.00 (3)	0.97 (9)	0.85 (2)	6.28
P11	0.967 (14)	0.91 (17)	0.89 (16.5)	0.85 (17)	0.87 (17.5)	0.87 (18)	0.87 (17.5)	0.83 (18)	0.91 (17.5)	0.98 (17)	0.91 (17.5)	0.85 (18)	0.73 (17.5)	0.85 (18)	0.73 (17.5)	0.82 (5)	16.47
P12	0.764 (21)	0.74 (22)	0.72 (23)	0.76 (21)	0.63 (21.5)	0.66 (22.5)	0.63 (21.5)	0.63 (22.5)	0.88 (19)	0.80 (22.5)	0.88 (19)	0.81 (21.5)	0.57 (19)	0.54 (22.5)	0.57 (19)	0.51 (21.5)	18.38
P13	0.948 (18)	0.98 (5.5)	0.95 (15)	0.91 (9)	0.95 (16)	0.97 (6.5)	0.95 (16)	0.90 (13)	0.95 (16)	1.00 (7.5)	0.95 (16)	0.90 (10)	0.88 (14)	0.97 (6.5)	0.88 (14)	0.75 (13)	12.25
P14	0.764 (21)	0.74 (22)	0.74 (21)	0.76 (21)	0.66 (19.5)	0.66 (22.5)	0.66 (19.5)	0.63 (22.5)	0.80 (21)	0.80 (22.5)	0.80 (21)	0.81 (21.5)	0.54 (20.5)	0.54 (22.5)	0.54 (20.5)	0.51 (21.5)	21.25
P15	0.751 (24)	0.78 (19)	0.745 (22)	0.68 (24)	0.66 (19.5)	0.78 (20)	0.66 (19.5)	0.64 (20)	0.79 (24)	0.86 (20)	0.79 (24)	0.69 (24)	0.54 (20.5)	0.61 (20)	0.54 (20.5)	0.45 (24)	21.56
P16	0.980 (4.5)	0.97 (10)	1.00 (4)	0.90 (12.5)	1.00 (5.5)	0.93 (12)	1.00 (5.5)	0.90 (13)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.89 (13.5)	1.00 (4)	0.93 (11)	1.00 (4)	0.70 (16.5)	8.41
P17	1.000 (1)	0.97 (10)	1.00 (4)	0.93 (7)	1.00 (5.5)	0.93 (12)	1.00 (5.5)	1.00 (5)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.92 (7.5)	1.00 (4)	0.93 (11)	1.00 (4)	0.80 (9.5)	6.53
P18	0.980 (4.5)	0.97 (10)	1.00 (4)	0.90 (12.5)	1.00 (5.5)	0.93 (12)	1.00 (5.5)	0.90 (13)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.89 (13.5)	1.00 (4)	0.93 (11)	1.00 (4)	0.70 (16.5)	8.41
P19	0.980 (4.5)	0.97 (10)	1.00 (4)	0.90 (12.5)	1.00 (5.5)	0.93 (12)	1.00 (5.5)	0.90 (13)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.89 (13.5)	1.00 (4)	0.93 (11)	1.00 (4)	0.70 (16.5)	8.41
P20	0.980 (4.5)	0.97 (10)	1.00 (4)	0.90 (12.5)	1.00 (5.5)	0.93 (12)	1.00 (5.5)	0.90 (13)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.89 (13.5)	1.00 (4)	0.93 (11)	1.00 (4)	0.70 (16.5)	8.41
P21	0.971 (9.5)	0.98 (5.5)	1.00 (4)	0.93 (7)	1.00 (5.5)	0.97 (6.5)	1.00 (5.5)	1.00 (5)	1.00 (5.5)	1.00 (7.5)	1.00 (5.5)	0.92 (7.5)	1.00 (4)	0.97 (6.5)	1.00 (4)	0.80 (9.5)	6.16
P22	0.968 (12)	1.00 (2.5)	0.98 (11.5)	0.93 (7)	1.00 (5.5)	1.00 (3)	1.00 (5.5)	1.00 (5)	0.98 (12)	1.00 (7.5)	0.98 (12)	0.92 (7.5)	0.95 (12)	1.00 (3)	0.95 (12)	0.80 (9.5)	7.97
P23	0.764 (21)	0.74 (22)	0.76 (18.5)	0.76 (21)	0.63 (21.5)	0.66 (22.5)	0.63 (21.5)	0.63 (22.5)	0.81 (20)	0.80 (22.5)	0.81 (20)	0.81 (21.5)	0.51 (22)	0.54 (22.5)	0.51 (22)	0.51 (21.5)	21.41
P24	0.764 (21)	0.74 (22)	0.76 (18.5)	0.76 (21)	1.00 (5.5)	1.00 (3)	1.00 (5.5)	1.00 (5)	0.98 (12)	1.00 (7.5)	0.98 (12)	0.92 (7.5)	0.95 (12)	1.00 (3)	0.95 (12)	0.80 (9.5)	11.06

Table 8
Performance and ranking of CNS microarray database.

Pipeline	Accuracy				Sensitivity				Specificity				Jaccard				Overall Average Grade
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	
P1	0.833 (4.5)	0.60 (17.5)	0.65 (12.5)	0.53 (21.5)	0.767 (5.5)	0.71 (1)	0.48 (12)	0.11 (21.5)	0.83 (14.5)	0.87 (8)	0.74 (12)	0.63 (20.5)	0.68 (3.5)	0.64 (1)	0.33 (12)	0.10 (21.5)	11.81
P2	0.800 (11)	0.68 (6.5)	0.65 (12.5)	0.65 (7)	0.767 (5.5)	0.63 (6.5)	0.48 (12)	0.53 (6.5)	0.79 (18.5)	0.89 (5)	0.74 (12)	0.71 (6.5)	0.58 (17.5)	0.51 (13.5)	0.33 (12)	0.32 (6.5)	9.94
P3	0.833 (4.5)	0.73 (3)	0.65 (12.5)	0.73 (3)	0.400 (23.5)	0.67 (2)	0.48 (12)	0.63 (3)	0.68 (23.5)	0.91 (1.5)	0.74 (12)	0.77 (3)	0.27 (23.5)	0.55 (2)	0.33 (12)	0.46 (3)	9.00
P4	0.800 (15.5)	0.67 (11)	0.65 (12.5)	0.60 (13.5)	0.717 (14)	0.61 (12)	0.48 (12)	0.16 (14.5)	0.87 (8.5)	0.78 (12)	0.74 (12)	0.66 (13.5)	0.64 (13.5)	0.52 (9)	0.33 (12)	0.12 (13.5)	12.44
P5	0.800 (15.5)	0.67 (11)	0.65 (12.5)	0.60 (13.5)	0.717 (14)	0.61 (12)	0.48 (12)	0.11 (21.5)	0.87 (8.5)	0.78 (12)	0.74 (12)	0.63 (20.5)	0.64 (13.5)	0.52 (9)	0.33 (12)	0.10 (21.5)	13.81
P6	0.833 (4.5)	0.60 (17.5)	0.63 (24)	0.53 (21.5)	0.800 (3)	0.51 (17)	0.48 (12)	0.11 (21.5)	0.87 (8.5)	0.68 (17)	0.74 (12)	0.63 (20.5)	0.70 (1.5)	0.47 (17)	0.33 (12)	0.10 (21.5)	14.44
P7	0.833 (4.5)	0.60 (17.5)	0.65 (12.5)	0.53 (21.5)	0.767 (5.5)	0.51 (17)	0.48 (12)	0.12 (18)	0.83 (14.5)	0.68 (17)	0.74 (12)	0.65 (17)	0.68 (3.5)	0.47 (17)	0.33 (12)	0.11 (17)	13.66
P8	0.767 (20)	0.60 (17.5)	0.65 (12.5)	0.53 (21.5)	0.767 (5.5)	0.51 (17)	0.48 (12)	0.11 (21.5)	0.79 (18.5)	0.68 (17)	0.74 (12)	0.63 (20.5)	0.58 (17.5)	0.47 (17)	0.33 (12)	0.10 (21.5)	16.47
P9	0.717 (22)	0.58 (21.5)	0.65 (12.5)	0.62 (8.5)	0.400 (23.5)	0.22 (24)	0.48 (12)	0.38 (10)	0.68 (23.5)	0.67 (20)	0.74 (12)	0.69 (9)	0.27 (23.5)	0.13 (24)	0.33 (12)	0.28 (9)	16.69
P10	0.800 (11)	0.60 (17.5)	0.65 (12.5)	0.53 (21.5)	0.817 (1.5)	0.51 (17)	0.48 (12)	0.11 (21.5)	0.85 (12)	0.68 (17)	0.74 (12)	0.63 (20.5)	0.70 (1.5)	0.47 (17)	0.33 (12)	0.10 (21.5)	14.25
P11	0.817 (9)	0.60 (17.5)	0.65 (12.5)	0.53 (21.5)	0.817 (1.5)	0.51 (17)	0.48 (12)	0.11 (21.5)	0.81 (16)	0.68 (17)	0.74 (12)	0.63 (20.5)	0.65 (9.5)	0.47 (17)	0.33 (12)	0.10 (21.5)	14.88
P12	0.717 (22)	0.58 (21.5)	0.65 (12.5)	0.62 (8.5)	0.650 (19)	0.23 (23)	0.48 (12)	0.16 (14.5)	0.72 (21)	0.64 (23)	0.74 (12)	0.66 (13.5)	0.38 (20)	0.14 (23)	0.33 (12)	0.12 (13.5)	15.13
P13	0.800 (15.5)	0.67 (11)	0.65 (12.5)	0.60 (13.5)	0.717 (14)	0.61 (12)	0.48 (12)	0.16 (14.5)	0.87 (8.5)	0.78 (12)	0.74 (12)	0.66 (13.5)	0.64 (13.5)	0.52 (9)	0.33 (12)	0.12 (13.5)	12.44
P14	0.717 (22)	0.52 (24)	0.65 (12.5)	0.60 (13.5)	0.467 (21.5)	0.26 (20.5)	0.48 (12)	0.40 (9)	0.79 (20)	0.60 (24)	0.74 (12)	0.66 (10)	0.35 (21)	0.16 (22)	0.33 (12)	0.24 (10)	14.06
P15	0.667 (24)	0.57 (23)	0.65 (12.5)	0.57 (18)	0.567 (20)	0.25 (22)	0.48 (12)	0.18 (11)	0.70 (22)	0.66 (21)	0.74 (12)	0.62 (24)	0.33 (22)	0.17 (21)	0.33 (12)	0.10 (18)	18.41
P16	0.833 (4.5)	0.73 (3)	0.65 (12.5)	0.73 (3)	0.733 (9.5)	0.67 (4)	0.48 (12)	0.63 (3)	0.90 (3.5)	0.89 (5)	0.74 (12)	0.77 (3)	0.67 (6.5)	0.55 (4)	0.33 (12)	0.46 (3)	6.28
P17	0.833 (4.5)	0.73 (3)	0.65 (12.5)	0.73 (3)	0.733 (9.5)	0.67 (4)	0.48 (12)	0.63 (3)	0.90 (3.5)	0.89 (5)	0.74 (12)	0.77 (3)	0.67 (6.5)	0.55 (4)	0.33 (12)	0.46 (3)	6.28
P18	0.833 (4.5)	0.73 (3)	0.65 (12.5)	0.73 (3)	0.733 (9.5)	0.62 (9)	0.48 (12)	0.63 (3)	0.90 (3.5)	0.92 (1.5)	0.74 (12)	0.77 (3)	0.67 (6.5)	0.51 (12)	0.33 (12)	0.46 (3)	6.88
P19	0.833 (4.5)	0.73 (3)	0.65 (12.5)	0.73 (3)	0.733 (9.5)	0.67 (4)	0.48 (12)	0.63 (3)	0.90 (3.5)	0.89 (5)	0.74 (12)	0.77 (3)	0.67 (6.5)	0.55 (4)	0.33 (12)	0.46 (3)	6.28
P20	0.800 (11)	0.68 (6.5)	0.67 (1)	0.70 (6)	0.708 (18)	0.63 (6.5)	0.48 (12)	0.53 (6.5)	0.94 (1)	0.89 (5)	0.74 (12)	0.71 (6.5)	0.65 (9.5)	0.51 (13.5)	0.33 (12)	0.32 (6.5)	8.34
P21	0.783 (19)	0.67 (11)	0.65 (12.5)	0.60 (13.5)	0.717 (14)	0.61 (12)	0.48 (12)	0.16 (14.5)	0.84 (13)	0.78 (12)	0.74 (12)	0.66 (13.5)	0.64 (13.5)	0.52 (9)	0.33 (12)	0.12 (13.5)	12.94
P22	0.800 (15.5)	0.67 (11)	0.65 (12.5)	0.60 (13.5)	0.717 (14)	0.62 (8)	0.48 (12)	0.16 (14.5)	0.87 (8.5)	0.81 (9)	0.74 (12)	0.66 (13.5)	0.64 (13.5)	0.53 (6)	0.33 (12)	0.12 (13.5)	11.81
P23	0.800 (15.5)	0.67 (11)	0.65 (12.5)	0.60 (13.5)	0.467 (21.5)	0.26 (20.5)	0.48 (12)	0.43 (8)	0.80 (17)	0.65 (22)	0.74 (12)	0.69 (8)	0.43 (19)	0.18 (20)	0.33 (12)	0.29 (8)	14.53
P24	0.800 (15.5)	0.67 (11)	0.65 (12.5)	0.60 (13.5)	0.717 (14)	0.61 (12)	0.48 (12)	0.16 (14.5)	0.87 (8.5)	0.78 (12)	0.74 (12)	0.66 (13.5)	0.64 (13.5)	0.52 (9)	0.33 (12)	0.12 (13.5)	12.44

Table 9
Performance and ranking of Lungcancer1 microarray database.

Pipeline	Accuracy				Sensitivity				Specificity				Jaccard				Overall Average Grade
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	
P1	0.978 (11)	1.000 (4)	0.990 (2)	0.938 (6.5)	1.00 (7)	1.00 (4)	0.99 (6)	1.00 (6)	0.80 (11)	0.90 (4)	0.80 (6)	0.67 (5.5)	0.976 (11)	1.00 (4)	0.99 (2)	0.93 (6.5)	6.03
P2	0.980 (8)	0.990 (10.5)	0.970 (11.5)	0.928 (10)	1.00 (7)	0.99 (10.5)	0.97 (11.5)	1.00 (6)	0.83 (8)	0.80 (10.5)	0.60 (11.5)	0.62 (9)	0.978 (8)	0.99 (10.5)	0.97 (11.5)	0.92 (16)	10.00
P3	0.98 (8)	0.990 (10.5)	0.981 (4.5)	0.928 (10)	1.00 (7)	0.99 (10.5)	1.00 (2.5)	1.00 (6)	0.83 (8)	0.80 (10.5)	0.83 (2.5)	0.62 (9)	0.978 (8)	0.99 (10.5)	0.98 (7.5)	0.92 (16)	8.19
P4	0.937 (16.5)	0.926 (16.5)	0.928 (18)	0.926 (14.5)	0.95 (16.5)	0.96 (16.5)	0.93 (18)	0.93 (16.5)	0.35 (16.5)	0.40 (16.5)	0.20 (18)	0.20 (16.5)	0.937 (16.5)	0.92 (16.5)	0.93 (18)	0.93 (11.5)	16.44
P5	0.937 (16.5)	0.926 (16.5)	0.970 (11.5)	0.926 (14.5)	0.95 (16.5)	0.96 (16.5)	0.97 (11.5)	0.93 (16.5)	0.35 (16.5)	0.40 (16.5)	0.60 (11.5)	0.20 (16.5)	0.937 (16.5)	0.92 (16.5)	0.97 (11.5)	0.93 (11.5)	14.81
P6	0.990 (3.5)	1.000 (4)	0.979 (8.5)	0.936 (8)	1.00 (7)	1.00 (4)	0.98 (8.5)	0.93 (16.5)	0.85 (3.5)	0.90 (4)	0.70 (8.5)	0.70 (3)	0.989 (3.5)	1.00 (4)	0.98 (4.5)	0.93 (11.5)	6.41
P7	0.990 (3.5)	1.000 (4)	0.969 (13)	0.980 (1)	1.00 (7)	1.00 (4)	0.97 (11.5)	0.93 (16.5)	0.85 (3.5)	0.90 (4)	0.60 (11.5)	0.83 (1)	0.989 (3.5)	1.00 (4)	0.97 (11.5)	0.98 (1)	6.28
P8	0.990 (3.5)	1.000 (4)	0.980 (6.5)	0.958 (4.5)	1.00 (7)	1.00 (4)	1.00 (2.5)	0.93 (16.5)	0.85 (3.5)	0.90 (4)	0.83 (2.5)	0.70 (3)	0.989 (3.5)	1.00 (4)	0.98 (7.5)	0.95 (4.5)	5.06
P9	0.894 (22.5)	0.894 (21.5)	0.905 (21)	0.905 (20.5)	0.90 (22.5)	0.90 (21.5)	0.91 (22.5)	0.91 (22.5)	0.0 (22)	0.00 (22)	0.00 (22.5)	0.00 (22.5)	0.894 (22.5)	0.89 (21.5)	0.91 (22.5)	0.91 (19.5)	21.84
P10	0.990 (3.5)	1.000 (4)	0.980 (6.5)	0.895 (23)	1.00 (7)	1.00 (4)	1.00 (2.5)	1.00 (6)	0.85 (3.5)	0.90 (4)	0.83 (2.5)	0.58 (13)	0.989 (3.5)	1.00 (4)	0.98 (7.5)	0.88 (23)	7.34
P11	0.978 (11)	1.000 (4)	0.990 (2)	0.938 (6.5)	1.00 (7)	1.00 (4)	0.99 (6)	1.00 (6)	0.80 (11)	0.90 (4)	0.80 (6)	0.67 (5.5)	0.976 (11)	1.00 (4)	0.99 (2)	0.93 (6.5)	6.03
P12	0.894 (22.5)	0.894 (21.5)	0.90 (23.5)	0.905 (20.5)	0.90 (22.5)	0.90 (21.5)	0.91 (22.5)	0.91 (22.5)	0.00 (22)	0.00 (22)	0.00 (22.5)	0.00 (22.5)	0.894 (22.5)	0.89 (21.5)	0.91 (22.5)	0.91 (19.5)	22.00
P13	0.937 (16.5)	0.926 (16.5)	0.937 (17)	0.926 (14.5)	0.95 (16.5)	0.96 (16.5)	0.94 (17)	0.93 (16.5)	0.35 (16.5)	0.40 (16.5)	0.30 (17)	0.20 (16.5)	0.937 (16.5)	0.92 (16.5)	0.94 (17)	0.93 (11.5)	16.19
P14	0.894 (22.5)	0.894 (21.5)	0.905 (23.5)	0.905 (20.5)	0.90 (22.5)	0.90 (21.5)	0.91 (22.5)	0.91 (22.5)	0.00 (22)	0.00 (22)	0.00 (22.5)	0.00 (22.5)	0.894 (22.5)	0.89 (21.5)	0.91 (22.5)	0.91 (19.5)	22.00
P15	0.894 (22.5)	0.894 (21.5)	0.905 (22)	0.905 (20.5)	0.90 (22.5)	0.90 (21.5)	0.91 (22.5)	0.91 (22.5)	0.00 (22)	0.00 (22)	0.00 (22.5)	0.00 (22.5)	0.894 (22.5)	0.89 (21.5)	0.91 (22.5)	0.91 (19.5)	21.91
P16	0.990 (3.5)	0.990 (10.5)	0.971 (10)	0.969 (2.5)	1.00 (7)	0.99 (10.5)	0.97 (11.5)	0.93 (16.5)	0.85 (3.5)	0.80 (10.5)	0.6 (11.5)	0.60 (11.5)	0.989 (3.5)	0.99 (10.5)	0.97 (11.5)	0.97 (2.5)	8.56
P17	0.967 (13)	0.990 (10.5)	0.979 (8.5)	0.914 (18)	1.00 (7)	0.99 (10.5)	0.98 (8.5)	0.93 (16.5)	0.78 (13)	0.80 (10.5)	0.70 (8.5)	0.63 (7)	0.964 (13)	0.99 (10.5)	0.98 (4.5)	0.91 (22)	11.34
P18	0.978 (11)	1.000 (4)	0.990 (2)	0.958 (4.5)	1.00 (7)	1.00 (4)	0.99 (6)	0.93 (16.5)	0.80 (11)	0.90 (4)	0.80 (6)	0.70 (3)	0.976 (11)	1.00 (4)	0.99 (2)	0.95 (4.5)	6.28
P19	0.990 (3.5)	0.990 (10.5)	0.960 (14)	0.969 (2.5)	1.00 (7)	0.99 (10.5)	0.96 (14)	0.97 (12.5)	0.85 (3.5)	0.80 (10.5)	0.50 (14)	0.60 (11.5)	0.989 (3.5)	0.99 (10.5)	0.96 (14)	0.97 (2.5)	9.03
P20	0.980 (8)	0.990 (10.5)	0.981 (4.5)	0.928 (10)	1.00 (7)	0.9 (10.5)	1.00 (2.5)	1.00 (6)	0.83 (8)	0.80 (10.5)	0.83 (2.5)	0.62 (9)	0.978 (8)	0.99 (10.5)	0.98 (7.5)	0.92 (16)	8.19
P21	0.937 (16.5)	0.926 (16.5)	0.947 (16)	0.926 (14.5)	0.95 (16.5)	0.96 (16.5)	0.95 (15.5)	0.93 (16.5)	0.35 (16.5)	0.40 (16.5)	0.40 (15.5)	0.20 (16.5)	0.937 (16.5)	0.92 (16.5)	0.95 (15.5)	0.93 (11.5)	15.84
P22	0.937 (16.5)	0.926 (16.5)	0.916 (19.5)	0.926 (14.5)	0.95 (16.5)	0.96 (16.5)	0.92 (19.5)	0.93 (16.5)	0.35 (16.5)	0.40 (16.5)	0.10 (19.5)	0.20 (16.5)	0.937 (16.5)	0.92 (16.6)	0.92 (19.5)	0.93 (11.5)	16.82
P23	0.905 (20)	0.851 (24)	0.916 (19.5)	0.873 (24)	0.91 (20)	0.90 (24)	0.92 (19.5)	0.92 (20)	0.00 (22)	0.00 (22)	0.10 (19.5)	0.03 (20)	0.905 (20)	0.85 (24)	0.92 (19.5)	0.87 (24)	21.38
P24	0.937 (16.5)	0.926 (16.5)	0.947 (15)	0.926 (14.5)	0.95 (16.5)	0.96 (16.5)	0.95 (15.5)	0.93 (16.5)	0.35 (16.5)	0.40 (16.5)	0.40 (15.5)	0.20 (16.5)	0.937 (16.5)	0.92 (16.5)	0.95 (15.5)	0.93 (11.5)	11.91

Table 10
Performance and ranking of Lungcancer2 microarray database.

Pipeline	Accuracy				Sensitivity				Specificity				Jaccard				Overall Average Grade
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	
P1	0.986 (4)	0.980 (5)	0.908 (12)	0.920 (11)	1.00 (4.5)	0.987 (4)	0.92 (15)	1.00 (4.5)	0.92 (4)	0.95 (3)	0.15 (13.5)	0.63 (8)	0.985 (4)	0.98 (6)	0.908 (10.5)	0.911 (22)	8.19
P2	0.992 (1.5)	0.979 (7.5)	0.979 (2)	0.931 (6.5)	1.00 (4.5)	0.986 (10.5)	0.98 (3)	1.00 (4.5)	0.95 (1.5)	0.90 (6.5)	0.80 (2.5)	0.65 (5.5)	0.992 (1.5)	0.98 (6)	0.979 (2)	0.925 (6.5)	4.50
P3	0.972 (10)	0.947 (13)	0.945 (5)	0.979 (1.5)	0.99 (10.5)	0.978 (12)	0.95 (5)	0.99 (9.5)	0.72 (11)	0.74 (12)	0.45 (5)	0.82 (1.5)	0.971 (10)	0.94 (12)	0.944 (5)	0.978 (1.5)	7.78
P4	0.881 (21.5)	0.916 (16.5)	0.908 (12)	0.914 (15.5)	0.92 (21.5)	0.932 (15.5)	0.92 (15)	0.91 (21.5)	0.20 (21.5)	0.40 (15.5)	0.15 (13.5)	0.20 (21)	0.881 (21.5)	0.92 (15.5)	0.908 (10.5)	0.914 (12.5)	14.97
P5	0.881 (21.5)	0.916 (16.5)	0.915 (8)	0.914 (15.5)	0.92 (21.5)	0.932 (15.5)	0.93 (8)	0.91 (21.5)	0.20 (21.5)	0.40 (15.5)	0.20 (9.5)	0.20 (21)	0.881 (21.5)	0.92 (15.5)	0.907 (14)	0.914 (12.5)	16.19
P6	0.986 (4)	0.980 (5)	0.908 (12)	0.920 (11)	1.00 (4.5)	0.987 (4)	0.92 (15)	1.00 (4.5)	0.92 (4)	0.95 (3)	0.15 (13.5)	0.63 (8)	0.985 (4)	0.98 (6)	0.908 (10.5)	0.911 (22)	8.19
P7	0.979 (7)	0.992 (1)	0.900 (19)	0.884 (24)	1.00 (4.5)	1 (1)	0.92 (15)	1.00 (4.5)	0.88 (7)	0.95 (3)	0.10 (19.5)	0.57 (10)	0.977 (8)	1.00 (1)	0.900 (18.5)	0.873 (24)	10.44
P8	0.979 (7)	0.986 (2.5)	0.893 (22.5)	0.927 (8.5)	1.00 (4.5)	0.993 (2)	0.91 (22.5)	1.00 (4.5)	0.88 (7)	0.95 (3)	0.05 (23.5)	0.70 (3.5)	0.977 (8)	0.99 (2)	0.893 (22.5)	0.919 (8.5)	9.50
P9	0.920 (16)	0.900 (22)	0.900 (19)	0.913 (21)	0.94 (16)	0.918 (21.5)	0.92 (15)	0.92 (16)	0.50 (16)	0.25 (21.5)	0.10 (19.5)	0.35 (16)	0.918 (16)	0.90 (21.5)	0.900 (18.5)	0.912 (18)	18.34
P10	0.979 (7)	0.986 (2.5)	0.893 (22.5)	0.927 (8.5)	1.00 (4.5)	0.918 (21.5)	0.91 (22.5)	1.00 (4.5)	0.88 (7)	0.25 (21.5)	0.05 (23.5)	0.70 (3.5)	0.977 (8)	0.90 (21.5)	0.893 (22.5)	0.919 (8.5)	13.09
P11	0.986 (4)	0.980 (5)	0.920 (7)	0.920 (11)	1.00 (4.5)	0.987 (4)	0.92 (15)	1.00 (4.5)	0.92 (4)	0.95 (3)	0.10 (19.5)	0.63 (8)	0.985 (4)	0.98 (6)	0.900 (18.5)	0.911 (22)	8.75
P12	0.920 (16)	0.900 (22)	0.913 (9)	0.913 (21)	0.94 (16)	0.918 (21.5)	0.93 (7)	0.92 (16)	0.50 (16)	0.25 (21.5)	0.30 (7.5)	0.35 (16)	0.918 (16)	0.90 (21.5)	0.912 (7)	0.912 (18)	15.75
P13	0.881 (21.5)	0.916 (16.5)	0.908 (12)	0.914 (15.5)	0.92 (21.5)	0.932 (15.5)	0.92 (15)	0.91 (21.5)	0.20 (21.5)	0.40 (15.5)	0.15 (13.5)	0.20 (21)	0.881 (21.5)	0.92 (15.5)	0.908 (10.5)	0.914 (12.5)	16.91
P14	0.920 (16)	0.900 (22)	0.900 (19)	0.913 (21)	0.94 (16)	0.918 (21.5)	0.92 (15)	0.92 (16)	0.50 (16)	0.25 (21.5)	0.10 (19.5)	0.35 (16)	0.918 (16)	0.90 (21.5)	0.900 (18.5)	0.912 (18)	15.59
P15	0.920 (16)	0.900 (22)	0.900 (19)	0.913 (21)	0.94 (16)	0.918 (21.5)	0.92 (15)	0.92 (16)	0.50 (16)	0.25 (21.5)	0.10 (19.5)	0.35 (16)	0.918 (16)	0.90 (21.5)	0.900 (18.5)	0.912 (18)	18.34
P16	0.960 (12)	0.966 (11.5)	0.947 (4)	0.947 (4)	0.97 (12.5)	0.986 (8.5)	0.9 (4)	0.95 (12.5)	0.70 (12.5)	0.78 (10.5)	0.45 (5)	0.50 (12.5)	0.959 (11.5)	0.96 (10.5)	0.946 (4)	0.947 (3.5)	8.69
P17	0.960 (12)	0.973 (10)	0.908 (15)	0.947 (4)	0.99 (10.5)	0.986 (8.5)	0.92 (15)	0.96 (11)	0.75 (10)	0.82 (8.5)	0.15 (13.5)	0.55 (11)	0.958 (13)	0.97 (8.5)	0.908 (10.5)	0.946 (5)	10.38
P18	0.979 (9)	0.978 (9)	0.978 (3)	0.979 (1.5)	0.99 (10.5)	0.986 (8.5)	0.99 (2)	0.99 (9.5)	0.82 (9)	0.82 (8.5)	0.80 (2.5)	0.82 (1.5)	0.978 (6)	0.97 (8.5)	0.979 (3)	0.978 (1.5)	5.84
P19	0.960 (12)	0.966 (11.5)	0.933 (6)	0.947 (4)	0.97 (12.5)	0.986 (8.5)	0.95 (5)	0.95 (12.5)	0.70 (12.5)	0.78 (10.5)	0.45 (5)	0.50 (12.5)	0.959 (11.5)	0.96 (10.5)	0.933 (6)	0.947 (6.5)	9.19
P20	0.992 (1.5)	0.979 (7.5)	0.986 (1)	0.931 (6.5)	1.00 (4.5)	0.986 (10.5)	1.00 (1)	1.00 (4.5)	0.95 (1.5)	0.90 (6.5)	0.90 (1)	0.65 (5.5)	0.992 (1.5)	0.98 (3.5)	0.986 (1)	0.925 (6.5)	4.00
P21	0.881 (21.5)	0.916 (16.5)	0.908 (12)	0.914 (15.5)	0.92 (21.5)	0.932 (15.5)	0.92 (15)	0.91 (21.5)	0.20 (21.5)	0.40 (15.5)	0.15 (13.5)	0.20 (21)	0.881 (21.5)	0.92 (15.5)	0.908 (10.5)	0.914 (12.5)	15.34
P22	0.881 (21.5)	0.916 (16.5)	0.846 (24)	0.914 (15.5)	0.92 (21.5)	0.932 (15.5)	0.85 (24)	0.91 (21.5)	0.20 (21.5)	0.40 (15.5)	0.30 (7.5)	0.20 (21)	0.881 (21.5)	0.92 (15.5)	0.846 (24)	0.914 (12.5)	17.13
P23	0.920 (16)	0.900 (22)	0.900 (19)	0.913 (21)	0.94 (16)	0.918 (21.5)	0.92 (15)	0.92 (16)	0.50 (16)	0.25 (21.5)	0.10 (19.5)	0.35 (16)	0.918 (16)	0.90 (21.5)	0.900 (18.5)	0.912 (18)	18.34
P24	0.881 (21.5)	0.916 (16.5)	0.907 (16)	0.914 (15.4)	0.92 (21.5)	0.932 (15.5)	0.92 (15)	0.91 (21.5)	0.20 (21.5)	0.40 (15.5)	0.20 (9.5)	0.20 (21)	0.881 (21.5)	0.92 (15.5)	0.906 (15)	0.914 (12.5)	17.18

Table 11
Performance and ranking of DLBCL Stanford microarray database.

Pipeline	Accuracy				Sensitivity				Specificity				Jaccard				Overall Average Grade
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	
P1	0.88 (7)	0.81 (12)	0.81 (17)	0.83 (15.5)	0.961 (8.5)	0.86 (12)	1.00 (9)	1.00 (6)	0.72 (9)	0.52 (12)	0.57 (17)	0.67 (7.5)	0.85 (7)	0.80 (11.5)	0.76 (17)	0.775 (15.5)	11.72
P2	0.94 (3)	0.81 (12)	0.87 (4)	0.83 (15.5)	0.986 (3)	0.86 (12)	1.00 (9)	1.00 (6)	0.88 (3)	0.52 (12)	0.73 (3.5)	0.67 (7.5)	0.92 (3)	0.80 (11.5)	0.83 (5.5)	0.775 (15.5)	8.13
P3	0.94 (3)	0.81 (12)	0.86 (7)	0.83 (15.5)	0.986 (3)	0.86 (12)	0.96 (20)	1.00 (6)	0.88 (3)	0.52 (12)	0.65 (9)	0.67 (7.5)	0.92 (3)	0.80 (11.5)	0.82 (11)	0.775 (15.5)	9.69
P4	0.81 (17.5)	0.81 (12)	0.86 (7)	0.84 (8.5)	0.865 (16.5)	0.86 (12)	0.96 (20)	0.83 (16.5)	0.52 (16.5)	0.52 (12)	0.65 (9)	0.60 (16.5)	0.80 (16.5)	0.80 (11.5)	0.82 (11)	0.829 (4.5)	12.66
P5	0.81 (17.5)	0.81 (12)	0.85 (10.5)	0.84 (8.5)	0.865 (16.5)	0.86 (12)	1.00 (9)	0.83 (16.5)	0.52 (16.5)	0.52 (12)	0.60 (14)	0.60 (16.5)	0.80 (16.5)	0.80 (11.5)	0.82 (11)	0.829 (4.5)	12.50
P6	0.88 (7)	0.81 (12)	0.84 (12)	0.83 (15.5)	0.961 (8.5)	0.86 (12)	1.00 (9)	1.00 (6)	0.72 (9)	0.52 (12)	0.67 (6)	0.67 (7.5)	0.85 (7)	0.80 (11.5)	0.80 (14)	0.775 (15.5)	9.47
P7	0.86 (12)	0.81 (12)	0.86 (7)	0.87 (2)	0.961 (8.5)	0.86 (12)	0.96 (20)	1.00 (6)	0.65 (12)	0.52 (12)	0.65 (9)	0.68 (2.5)	0.82 (12)	0.80 (11.5)	0.82 (11)	0.825 (8.5)	10.44
P8	0.86 (12)	0.81 (12)	0.86 (7)	0.87 (2)	0.961 (8.5)	0.86 (12)	0.96 (20)	1.00 (6)	0.65 (12)	0.52 (12)	0.65 (9)	0.68 (2.5)	0.82 (12)	0.80 (11.5)	0.82 (11)	0.825 (8.5)	10.44
P9	0.80 (22.5)	0.81 (12)	0.76 (22.5)	0.77 (22)	0.857 (22)	0.86 (12)	1.00 (9)	0.80 (22)	0.65 (12)	0.52 (12)	0.17 (24)	0.40 (22)	0.78 (22)	0.80 (11.5)	0.74 (23)	0.755 (22)	17.63
P10	0.86 (12)	0.81 (12)	0.83 (14.5)	0.87 (2)	0.961 (8.5)	0.86 (12)	1.00 (9)	1.00 (6)	0.65 (12)	0.52 (12)	0.73 (3.5)	0.70 (1)	0.82 (12)	0.80 (11.5)	0.83 (5.5)	0.833 (1)	9.06
P11	0.88 (7)	0.81 (12)	0.83 (14.5)	0.83 (15.5)	0.961 (8.5)	0.86 (12)	1.00 (9)	1.00 (6)	0.72 (9)	0.52 (12)	0.69 (5)	0.67 (7.5)	0.85 (7)	0.80 (11.5)	0.78 (16)	0.775 (15.5)	10.75
P12	0.80 (22.5)	0.81 (12)	0.76 (22.5)	0.77 (22)	0.857 (22)	0.86 (12)	1.00 (9)	0.80 (22)	0.47 (22)	0.52 (12)	0.24 (21.5)	0.40 (22)	0.78 (22)	0.80 (11.5)	0.75 (21.5)	0.755 (22)	18.00
P13	0.81 (17.5)	0.81 (12)	0.84 (13)	0.84 (8.5)	0.865 (16.5)	0.86 (12)	0.83 (24)	0.83 (16.5)	0.52 (16.5)	0.52 (12)	0.60 (14)	0.60 (16.5)	0.80 (16.5)	0.80 (11.5)	0.83 (4)	0.829 (4.5)	13.16
P14	0.80 (22.5)	0.81 (12)	0.78 (19)	0.77 (22)	0.857 (22)	0.86 (12)	1.00 (9)	0.80 (22)	0.47 (22)	0.52 (12)	0.30 (19)	0.40 (22)	0.78 (22)	0.80 (11.5)	0.76 (19)	0.755 (22)	17.47
P15	0.83 (14)	0.81 (12)	0.75 (24)	0.77 (22)	0.857 (22)	0.86 (12)	1.00 (9)	0.80 (22)	0.47 (22)	0.52 (12)	0.21 (23)	0.40 (22)	0.78 (22)	0.80 (11.5)	0.72 (24)	0.755 (22)	17.06
P16	0.86 (9.5)	0.81 (12)	0.78 (19)	0.85 (4.5)	0.907 (12.5)	0.86 (12)	1.00 (9)	0.94 (12.5)	0.78 (6.5)	0.52 (12)	0.30 (19)	0.66 (12.5)	0.84 (9.5)	0.80 (11.5)	0.76 (19)	0.818 (10.5)	11.91
P17	0.94 (3)	0.81 (12)	0.88 (3)	0.83 (15.5)	0.986 (3)	0.86 (12)	1.00 (9)	1.00 (6)	0.88 (3)	0.52 (12)	0.63 (12)	0.67 (7.5)	0.92 (3)	0.80 (11.5)	0.85 (2)	0.775 (15.5)	8.38
P18	0.94 (3)	0.81 (12)	0.89 (2)	0.83 (15.5)	0.986 (3)	0.86 (12)	1.00 (9)	1.00 (6)	0.88 (3)	0.52 (12)	0.80 (2)	0.67 (7.5)	0.92 (3)	0.80 (11.5)	0.85 (3)	0.775 (15.5)	7.75
P19	0.86 (9.5)	0.81 (12)	0.86 (7)	0.85 (4.5)	0.907 (12.5)	0.86 (12)	0.96 (20)	0.94 (12.5)	0.78 (6.5)	0.52 (12)	0.65 (9)	0.66 (2.5)	0.84 (9.5)	0.80 (11.5)	0.82 (11)	0.818 (10.5)	10.72
P20	0.94 (3)	0.85 (1)	0.90 (1)	0.83 (15.5)	0.986 (3)	0.96 (1)	1.00 (9)	1.00 (6)	0.88 (3)	0.65 (1)	0.82 (7.5)	0.67 (3)	0.92 (3)	0.81 (1)	0.87 (1)	0.775 (15.5)	4.13
P21	0.81 (17.5)	0.81 (12)	0.8(16(16)	0.84 (8.5)	0.865 (16.5)	0.86 (12)	0.90 (23)	0.83 (16.5)	0.52 (16.5)	0.52 (12)	0.58 (16)	0.60 (16.5)	0.80 (16.5)	0.80 (11.5)	0.79 (15)	0.829 (4.5)	13.09
P22	0.81 (17.5)	0.81 (12)	0.85 (10.5)	0.84 (8.5)	0.865 (16.5)	0.86 (12)	1.00 (9)	0.83 (16.5)	0.52 (16.5)	0.52 (12)	0.60 (14)	0.60 (16.5)	0.80 (16.5)	0.80 (11.5)	0.82 (11)	0.829 (4.5)	12.50
P23	0.80 (22.5)	0.68 (23.5)	0.76 (21)	0.77 (22)	0.857 (22)	0.76 (24)	1.00 (9)	0.80 (22)	0.47 (22)	0.16 (24)	0.24 (21.5)	0.40 (22)	0.78 (22)	0.67 (24)	0.75 (21.5)	0.755 (22)	21.69
P24	0.81 (17.5)	0.68 (23.5)	0.78 (19)	0.84 (8.5)	0.865 (16.5)	0.86 (12)	1.00 (9)	0.83 (16.5)	0.52 (16.5)	0.52 (12)	0.30 (19)	0.60 (16.5)	0.80 (16.5)	0.80 (11.5)	0.76 (19)	0.829 (4.5)	14.56

Table 12
Top average grade models for different microarray datasets.

Datasets	Pipelines with top3 grades		
	Top1	Top2	Top3
Leukemia	[P6]	[P1,P7,P11]	[P8]
Colon	[P22]	[P7]	[P24]
Prostate	[P8]	[P21]	[P7,P10]
CNS	[P16,P17,P19]	[P18]	[P20]
Lungcancer1	[P8]	[P1,P11]	[P7,P18]
Lungcancer2	[P20]	[P2]	[P18]
DLBCL Stanford	[P20]	[P18]	[P2]

observed data. Every value of the independent variable is associated with a value of the dependent variable [51,52].

4.7. Performance metrics

The performance metrics considered for evaluation of the experimental work are discussed here. In the final stage of classification the actual output is compared with target output and there are four possible outcomes.

- i True positive (TrP): Diseased sample correctly classified
- ii False positive (FlP): Diseased sample incorrectly classified
- iii True negative (TrN): Healthy sample correctly identified
- iv False negative (FlN): Healthy sample incorrectly classified

Based on these outcomes four performance criteria are evaluated

$$\text{Accuracy} = \left(\frac{\text{TrP} + \text{TrN}}{\text{Total number of sample}} \right) * 100 \quad (8)$$

$$\text{Sensitivity} = \left(\frac{\text{TrP}}{\text{TrP} + \text{FlP}} \right) * 100 \quad (9)$$

$$\text{Specificity} = \left(\frac{\text{TrN}}{\text{TrN} + \text{FlN}} \right) * 100 \quad (10)$$

$$\text{Jaccard} = \left(\frac{\text{TrP}}{\text{TrP} + \text{FlP} + \text{FlN}} \right) * 100 \quad (11)$$

4.8. Results and discussion

The results of four performance metrics evaluated for twenty four pipelines with four classifiers for seven microarray databases are shown in Tables 5–11.

The results of Leukemia microarray database for the four performance metrics/classifier for twenty four pipelines are shown in Table 5. In each cell the results are followed by the grade of the results in the respective column for twenty four pipelines. Highest performance of a pipeline is assigned grade 1, next highest as 2 and so on. When more than one pipeline has the same level of performance, average grade is assigned to each one. For example in the second column of Table 5, P10 yield the highest accuracy of 0.98 for KNN and it is assigned grade 1. Next highest accuracy of 0.95 for KNN is yield by P2, P3 and P20, their respective grades 2, 3, and 4, average of which i.e. $(2 + 3 + 4)/3 = 3$ is assigned as the grade to all the three pipeline models.

To evaluate the best model for each dataset, the last column of the table shows the overall average grade of each pipeline. The overall average grade of a pipeline model is the average of the grades of all the classifiers for all the performance metrics of the respective pipeline model.

From the analysis overall average grade of Table 5 for Leukemia database, it is revealed that model P6 possesses the lowest overall average grade of 5.22 and hence can be considered as the best model for Leukemia database. Further after P6, the models P1, P7 and P11

possess the next lowest overall average grade of 5.78 and can be considered as the next best models.

The overall average grade of for Colon database presented in Table 6 is analyzed and it is found that model P22 possesses the lowest overall average grade of 7.56 and hence may be treated as the best model for Colon database. Further after P22, the model P7 possesses the next lowest overall average grade of 7.91 and can be considered as the next best model.

From the analysis overall average grade of Table 7 for Prostate database, it is revealed that model P8 possesses the lowest overall average grade of 5.34 and hence can be considered as the best model for Prostate database. Further after P8, the model P21 possesses the next lowest overall average grade of 6.16 and can be considered as the next best models for Prostate database.

The overall average grade of for CNS database presented in Table 8 is analyzed and it is found that models P16, 17, and P19 possess the lowest overall average grade of 6.28 and hence may be treated as the best models for CNS database. Further after these models, the models P18 possesses the next lowest overall average grade of 6.88 and can be considered as the next best model.

From the analysis of overall average grade of Table 9 for Lungcancer1 database, it is revealed that model P8 possesses the lowest overall average grade of 5.06 and hence can be considered as the best model for Lungcancer1 database. Further after P8, the models P1 and P11 possess the next lowest overall average grade of 6.03 and can be considered as the next best models for Lungcancer1 database.

The overall average grade of for Lungcancer2 database presented in Table 10 is analyzed and it is found that model P20 possesses the lowest overall average grade of 4.00 and hence may be treated as the best model for Lungcancer2 database. Further after P20 model, the model P2 possesses the next lowest overall average grade of 4.5 and can be considered as the next best model.

From the analysis of overall average grade of Table 11 for DLBCL Stanford database, it is revealed that model P20 possesses the lowest overall average grade of 4.13 and hence can be considered as the best model for DLBCL Stanford database. Further after P20, the model P18 possesses the next lowest overall average grade of 7.75 and can be considered as the next best model for DLBCL Stanford database.

The pipeline models with the top three grades obtained from the overall average grade from Tables 5–11 is presented below in Table 12. From such average grading/pipeline, it is difficult to conclude the performance of any pipeline which is better in comparison to others when all the databases are considered simultaneously.

As the average results of four classifiers for each performance index are considered here for analysis. After 1st stage grading, the grades of Tables 5–11 for different databases are average out and is presented in Table 13. For example if the average grade of P1 for accuracy/KNN in Table 13 is referred as $T_{13}(P_1, \text{Accuracy}, \text{KNN})$. Then

$$T_{13}(P_1, \text{Accuracy}, \text{KNN}) = [T_5(P_1, \text{Accuracy}, \text{KNN}) + T_6(P_1, \text{Accuracy}, \text{KNN}) + \dots + T_{11}(P_1, \text{Accuracy}, \text{KNN})] / 7$$

where T_5, \dots, T_{11} are Tables 5–11, respectively.

It can be revealed that a single pipeline does not possess the highest grade or lowest grade for different classifiers and different performance metric. So obtaining a conclusive result after 1st stage grading is also difficult. Therefore to arrive at a consensus second stage of grading is done. Table 14 shows the second stage grading of the pipeline. For example taking average grade of accuracy for KNN classifier in Table 13, the lowest average grade is 5.43 of pipeline P1, which is assigned grade 1. The next lowest average grade is 5.64 of pipeline P2, which is assigned grade 2, and so on.

Table 13
Average of 1st stage grading for seven databases.

pipeline	Accuracy				Sensitivity				Specificity				Jaccard			
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR
P1	5.43	9.57	9.36	12.57	9.21	7.21	9.79	10.29	11.07	8.36	11.21	10.79	9.21	7.43	9.50	12.43
P2	5.64	10.00	9.29	9.00	7.00	10.79	9.14	7.14	8.43	9.79	8.93	7.57	9.07	10.50	8.29	11.00
P3	5.93	10.29	9.36	7.71	10.43	10.36	11.07	7.36	11.43	10.07	8.07	6.50	11.29	10.07	11.43	9.29
P4	15.64	12.93	11.86	12.00	16.58	12.57	13.71	14.21	15.50	12.50	12.21	13.93	16.50	12.57	12.21	11.07
P5	18.36	12.93	10.50	12.00	16.21	12.57	11.36	15.21	15.50	12.50	10.86	14.93	16.50	12.57	12.36	12.21
P6	8.36	9.57	14.43	12.79	9.64	9.50	11.36	11.79	9.79	9.64	12.79	10.43	8.64	9.71	10.86	12.64
P7	7.29	6.93	12.43	9.57	6.71	6.93	12.79	9.43	7.43	8.29	11.86	6.93	6.50	6.86	11.43	9.86
P8	9.50	7.14	12.29	8.29	6.71	7.07	12.71	9.93	8.00	8.29	11.86	6.79	8.50	7.00	12.29	8.79
P9	20.29	20.43	17.43	18.83	20.57	20.79	17.36	19.29	20.14	20.29	17.64	18.93	20.57	20.71	19.86	18.86
P10	7.14	7.14	13.21	11.79	6.14	9.86	12.00	10.07	7.07	10.93	11.21	9.64	6.21	9.79	11.64	12.00
P11	10.07	9.57	9.43	12.57	8.64	9.50	11.57	10.29	11.29	9.64	10.86	10.79	10.07	9.71	12.50	12.43
P12	20.64	20.43	18.36	18.83	20.79	20.64	14.29	19.93	20.86	20.71	15.00	19.57	20.57	20.57	15.71	19.08
P13	14.21	12.71	13.21	10.67	15.07	13.14	16.43	14.93	13.43	12.21	13.86	14.29	14.29	13.00	13.07	11.07
P14	20.29	20.79	18.57	19.67	20.86	20.29	17.71	19.14	20.64	20.86	18.36	19.07	20.93	20.43	19.79	18.50
P15	20.14	20.21	19.21	20.92	19.79	20.14	16.07	19.07	20.50	20.07	19.21	21.43	20.21	19.93	18.29	20.25
P16	9.14	10.07	10.36	5.79	9.07	10.14	8.36	10.21	6.86	10.36	12.57	9.50	7.21	10.14	9.57	7.29
P17	9.07	9.86	10.93	10.14	8.93	9.21	11.07	9.50	8.71	9.57	12.36	8.36	9.36	9.07	8.64	11.36
P18	7.64	8.29	7.79	9.14	8.00	9.00	8.21	12.21	6.86	8.14	7.93	7.86	8.50	9.29	6.71	9.86
P19	9.14	10.07	8.57	5.79	9.07	10.14	11.57	9.14	6.86	10.36	11.14	8.50	7.21	10.14	9.50	6.79
P20	6.21	8.43	4.79	8.86	8.00	9.21	7.29	7.14	6.43	8.21	6.07	7.57	7.43	9.00	7.21	10.50
P21	14.79	12.86	11.50	13.21	12.93	13.00	12.71	15.93	13.14	11.93	11.57	16.07	12.64	12.64	10.79	12.79
P22	14.64	12.00	14.64	12.36	13.93	11.50	14.71	14.21	13.43	11.50	12.93	14.36	14.79	11.29	14.86	11.07
P23	19.36	20.93	18.07	20.25	19.93	22.36	16.29	18.64	19.21	22.29	17.43	18.43	19.64	22.29	18.64	18.92
P24	17.07	16.86	14.43	14.43	14.43	12.50	10.71	13.64	14.43	11.93	12.36	15.29	15.14	12.14	13.14	10.86

Table 14
Second stage grading of the pipeline models for different performance metrics.

Pipeline	2nd stage grading for Accuracy				Average Grade	2nd stage grading for Sensitivity				Average Grade	2nd stage grading wrt Specificity				Average Grade	2nd stage grading wrt Jaccard				Average Grade
	KNN	NB	ANN	MLR		KNN	NB	ANN	MLR		KNN	NB	ANN	MLR		KNN	NB	ANN	MLR	
P1	1	7	5.5	15.5	7.25	11	3	5	10.5	4	11	5	8.5	12.5	9.25	10	3	5.5	16.5	8.75
P2	2	10	4	6	5.5	4	13	4	1.5	5.625	8	9	4	4.5	6.375	9	13	3	9	8.5
P3	3	13	5.5	3	6.125	13	12	9.5	3	4	13	10	3	1	6.75	13	10	10.5	4	9.375
P4	17	17.5	12	12.5	14.75	19	16.5	17	15.5	17	18.5	19	13	14	16	18.5	16.5	13	11	14.75
P5	19	17.5	9	12.5	14.5	18	16.5	7.5	18	4	18.5	19	5.5	17	14.875	18.5	16.5	15	15	16.25
P6	8	7	17.5	17	12.375	12	7.5	7.5	12	9.75	10	7.5	17	11	11.375	8	7.5	9	18	10.625
P7	6	1	14	8	7.25	2.5	1	16	5	4	6	3.5	11.5	3	6	2	1	10.5	5.5	4.75
P8	12	2.5	13	4	7.875	2.5	2	14.5	7	6.5	7	3.5	11.5	2	6	6.5	2	14	3	6.375
P9	22.5	21.5	20	20.5	21.125	22	23	23	23	4	21	21	22	21	21.25	22.5	23	24	21	22.625
P10	5	2.5	15.5	11	8.5	1	9	13	8	7.75	5	13	8.5	10	9.125	1	9	12	14	9
P11	13	7	7	15.5	10.625	7	7.5	11.5	10.5	4	12	7.5	5.5	12.5	9.375	12	7.5	16	16.5	13
P12	24	21.5	22	20.5	22	23	22	18	24	21.75	24	22	20	23	22.25	22.5	22	20	23	21.875
P13	14	15	15.5	10	13.625	17	19	22	17	4	15.5	17	19	15	16.625	15	19	17	11	15.5
P14	22.5	23	23	22	22.625	24	21	24	22	22.75	23	23	23	22	22.75	24	21	23	20	22
P15	21	20	24	24	22.25	20	20	20	21	4	22	20	24	24	22.5	21	20	21	24	21.5
P16	10.5	11.5	8	1.5	7.875	9.5	10.5	3	9	8	3	12	16	9	9.875	3.5	11.5	7	2	6
P17	9	9	10	9	9.25	8	5.5	9.5	6	4	9	6	14.5	7	9.125	11	5	4	13	8.25
P18	7	4	2	7	5	5.5	4	2	13	6.125	3	1	2	6	3	6.5	6	1	5.5	4.75
P19	10.5	11.5	3	1.5	6.625	9.5	10.5	11.5	4	4	3	12	7	8	7.375	3.5	11.5	5.5	1	5.375
P20	4	5	1	5	3.75	5.5	5.5	1	1.5	3.375	1	2	1	4.5	2.125	5	4	2	7	4.5
P21	16	16	11	18	15.25	14	18	14.5	19	4	14	16	10	19	14.625	14	18	8	19	14.75
P22	15	14	19	14	15.5	15	14	19	15.5	15.875	15.5	14	18	16	15.875	16	14	19	11	15
P23	20	24	21	23	22	21	24	21	20	4	20	24	21	20	21.25	20	24	22	22	22
P24	18	19	17.5	19	18.375	16	15	6	14	12.75	17	16	14.5	18	16.25	17	15	18	8	14.5

Table 15
Significance difference for four performance metrics.

Pipeline	Significance Difference (maximum average grade-average grade of specified pipeline)			
	Accuracy	Sensitivity	Specificity	Jaccard
P20	22.625 – 3.75 = 18.875	22.75 – 3.37 = 19.38	22.75 – 2.12 = 20.63	22.625 – 4.5 = 18.125
P18	22.625 – 5 = 17.625	22.75 – 6.125 = 16.625*	22.75 – 3 = 19.75	22.625 – 4.75 = 17.875
P2	22.625 – 5.5 = 17.125	22.75 – 5.6 = 17.15	22.75 – 6.375 = 16.375*	22.625 – 8.5 = 14.125*

The average of second stage grading for each pipelines in Table 14 is presented as the last column under each performance metric. It can be seen that the average grade of model P20 is 3.75 for accuracy, is 3.375 for sensitivity, is 2.125 for specificity and

4.5 for jaccard performance metrics which is the minimum, which indicates that P20 may be considered as the most stable model. Further in this proposed approach the performance of multiple models over multiple classifiers and on multiple data sets is com-

Table 16
Comparison of performance of P20 with single ranking techniques with their 1st order grading.

Database	Ranking methods	Accuracy				Sensitivity				Specificity				Jaccard			
		KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR
Leukemia	SNR	0.69 (5)	0.59 (5)	0.61 (5)	0.57 (5)	0.15 (5)	0.05 (4)	0.32 (5)	0.00 (5)	0.73 (5)	0.67 (5)	0.69 (5)	0.64 (5)	0.10 (5)	0.05 (5)	0.30 (5)	0.00 (5)
	PCC	0.98 (1.5)	1.00 (2)	0.89 (4)	0.98 (1.5)	0.90 (2)	1.00 (2)	0.78 (4)	0.90 (1.5)	0.98 (2.5)	1.00 (2.5)	0.98 (2.5)	0.98 (1.5)	0.90 (2)	1.00 (2)	0.78 (4)	0.90 (1.5)
	IG	0.98 (1.5)	1.00 (2)	0.94 (3)	0.98 (1.5)	0.90 (2)	1.00 (2)	0.85 (3)	0.90 (1.5)	0.98 (2.5)	1.00 (2.5)	0.98 (2.5)	0.98 (1.5)	0.90 (2)	1.00 (2)	0.85 (3)	0.90 (1.5)
	ttest	0.94 (4)	0.97 (4)	0.97 (2)	0.88 (4)	0.85 (4)	0.95 (3)	0.90 (2)	0.70 (3.5)	0.98 (2.5)	1.00 (2.5)	0.97 (4)	0.93 (3)	0.85 (4)	0.95 (4)	0.90 (2)	0.65 (4)
	P20	0.95 (3)	1.00 (2)	0.98 (1)	0.92 (3)	0.90 (1)	1.00 (2)	0.95 (1)	0.70 (3.5)	0.98 (2.5)	1.00 (2.5)	1.00 (1)	0.92 (1)	0.90 (2)	1.00 (2)	0.95 (1)	0.70 (3)
	SNR	0.61 (3)	0.54 (5)	0.56 (4.5)	0.61 (2)	0.37 (5)	0 (5)	0.5 (4)	0.35 (4.5)	0.7 (5)	0.60 (4.5)	0.6 (4.5)	0.64 (4.5)	0.26 (5)	0 (5)	0.33 (4)	0.22 (4.5)
Colon	PCC	0.74 (1)	0.57 (3)	0.57 (2.5)	0.05 (5)	0.58 (1)	0.72 (1.5)	0.62 (1.5)	0.35 (4.5)	0.77 (2)	0.60 (4.5)	0.72 (1.5)	0.64 (4.5)	0.47 (2)	0.54 (1.5)	0.47 (1.5)	0.22 (4.5)
	IG	0.57 (4.5)	0.57 (3)	0.56 (4.5)	0.06 (3.5)	0.38 (3.5)	0.72 (1.5)	0.10 (5)	0.65 (2)	0.78 (1)	0.78 (1.5)	0.60 (4.5)	0.75 (3)	0.27 (4)	0.54 (1.5)	0.03 (5)	0.43 (2.5)
	ttest	0.57 (4.5)	0.57 (3)	0.57 (2.5)	0.06 (3.5)	0.38 (3.5)	0.67 (3)	0.62 (1.5)	0.65 (2)	0.76 (3.5)	0.78 (1.5)	0.72 (1.5)	0.76 (2)	0.48 (1)	0.53 (3)	0.47 (1.5)	0.43 (2.5)
	P20	0.68 (2)	0.74 (1)	0.71 (1)	0.81 (1)	0.43 (2)	0.60 (4)	0.58 (3)	0.65 (2)	0.76 (3.5)	0.70 (3)	0.69 (3)	0.80 (1)	0.39 (3)	0.50 (4)	0.35 (3)	0.52 (1)
	SNR	0.57 (2.5)	0.57 (2.5)	0.57 (2.5)	0.46 (3.5)	0.65 (3.5)	0.63 (3)	0.45 (3)	0.21 (2)	0.45 (4)	0.45 (4)	0.45 (4)	0.14 (3)	0.63 (3)	0.63 (1)	0.33 (3)	0.28 (3.5)
CNS	PCC	0.57 (2.5)	0.57 (2.5)	0.57 (2.5)	0.46 (3.5)	0.65 (3.5)	0.63 (3)	0.40 (4.5)	0.10 (5)	0.40 (5)	0.40 (5)	0.45 (2)	0.92 (1)	0.40 (4)	0.33 (3)	0.28 (3.5)	
	IG	0.45 (4.5)	0.45 (4.5)	0.45 (4.5)	0.46 (3.5)	0.70 (2)	0.63 (3)	0.62 (1)	0.20 (3.5)	0.62 (3)	0.97 (1)	0.97 (1)	0.13 (4.5)	0.45 (4.5)	0.45 (3.5)	0.35 (1)	0.28 (3.5)
	ttest	0.45 (4.5)	0.45 (4.5)	0.45 (4.5)	0.46 (3.5)	0.64 (5)	0.63 (3)	0.40 (4.5)	0.20 (3.5)	0.80 (2)	0.80 (2)	0.80 (2)	0.13 (4.5)	0.45 (4.5)	0.45 (3.5)	0.30 (5)	0.28 (3.5)
	P20	0.80 (1)	0.68 (1)	0.67 (1)	0.70 (1)	0.71 (1)	0.63 (3)	0.48 (2)	0.53 (1)	0.94 (1)	0.89 (3)	0.74 (3)	0.71 (1)	0.65 (2)	0.51 (2)	0.33 (3)	0.32 (1)
	SNR	0.57 (2.5)	0.57 (2.5)	0.57 (2.5)	0.46 (3.5)	0.65 (3.5)	0.63 (3)	0.45 (3)	0.21 (2)	0.45 (4)	0.45 (4)	0.45 (4)	0.14 (3)	0.63 (3)	0.63 (1)	0.33 (3)	0.28 (3.5)

Table 16 (Continued)

Database	Ranking methods	Accuracy				Sensitivity				Specificity				Jaccard			
		KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR
Prostrate	SNR	0.53 (4)	0.05 (3.5)	0.05 (3.5)	0.52 (5)	0.72 (5)	0.44 (5)	0.20 (5)	0.58 (5)	0.82 (5)	0.69 (5)	0.60 (5)	0.72 (5)	0.54 (5)	0.30 (5)	0.12 (5)	0.42 (5)
	PCC	0.54 (3)	0.05 (3.5)	0.05 (3.5)	0.54 (3)	1.00 (2.5)	0.93 (3)	1.00 (2.5)	0.93 (3)	0.98 (3.5)	1.00 (2.5)	0.96 (3.5)	0.91 (2.5)	0.95 (3.5)	0.93 (3)	0.92 (4)	0.77 (2)
	IG	0.54 (3)	0.05 (3.5)	0.05 (3.5)	0.54 (3)	1.00 (2.5)	0.97 (1)	1.00 (2.5)	1.00 (1)	1.00 (1.5)	1.00 (2.5)	1.00 (1.5)	0.91 (2.5)	1.00 (1.5)	0.97 (1)	1.00 (1.5)	0.70 (3.5)
	ttest	0.54 (3)	0.05 (3.5)	0.05 (3.5)	0.54 (3)	1.00 (2.5)	0.93 (3)	1.00 (2.5)	0.97 (2)	0.98 (3.5)	1.00 (2.5)	0.96 (3.5)	0.98 (1)	0.95 (3.5)	0.93 (3)	0.90 (3)	0.90 (1)
	P20	0.98 (1)	0.97 (1)	1.00 (1)	0.90 (1)	1.00 (2.5)	0.93 (3)	1.00 (2.5)	0.90 (4)	1.00 (1.5)	1.00 (2.5)	1.00 (1.5)	0.89 (4)	1.00 (1.5)	0.93 (3)	1.00 (1.5)	0.70 (3.5)
Lungcancer1	SNR	0.88 (2.5)	0.88 (5)	0.72 (4.5)	0.86 (4)	1.00 (2)	1.00 (1)	0.80 (4.5)	0.94 (4)	0.90 (1.5)	0.90 (1)	0.21 (4.5)	0.33 (4)	1.00 (1)	1.00 (1)	0.33 (4.5)	0.85 (4)
	PCC	0.70 (5)	0.87 (3)	0.72 (4.5)	0.87 (2.5)	0.95 (5)	0.90 (2.5)	0.80 (4.5)	0.92 (5)	0.40 (5)	0.45 (5)	0.21 (4.5)	0.20 (3)	0.94 (4.5)	0.93 (5)	0.33 (4.5)	0.91 (3)
	IG	0.88 (2.5)	0.87 (3)	0.88 (2.5)	0.85 (5)	1.00 (2)	0.85 (4.5)	1.00 (2)	0.95 (3)	0.90 (1.5)	0.75 (4)	0.90 (1.5)	0.16 (5)	0.94 (4.5)	0.97 (4)	1.00 (1.5)	0.72 (5)
	ttest	0.87 (4)	0.87 (3)	0.88 (2.5)	0.87 (2.5)	0.99 (4)	0.85 (4.5)	1.00 (2)	1.00 (1.5)	0.80 (4)	0.80 (2.5)	0.90 (1.5)	0.65 (1)	0.99 (2)	0.99 (2.5)	1.00 (1.5)	0.94 (1)
	P20	0.98 (1)	0.99 (1)	0.98 (1)	0.93 (1)	1.00 (2)	0.90 (2.5)	1.00 (2)	1.00 (1.5)	0.83 (3)	0.80 (2.5)	0.83 (3)	0.62 (2)	0.98 (3)	0.99 (2.5)	0.98 (3)	0.92 (2)
Lungcancer2	SNR	0.91 (4)	0.93 (4)	0.90 (4.5)	0.89 (3)	0.93 (4)	0.94 (4)	0.90 (2.5)	0.91 (4.5)	0.77 (2.5)	0.45 (4)	0.00 (4.5)	0.05 (5)	0.91 (4)	0.93 (4)	0.90 (4.5)	0.89 (3)
	PCC	0.88 (5)	0.92 (5)	0.90 (4.5)	0.91 (2)	0.92 (5)	0.93 (5)	0.90 (2.5)	0.91 (4.5)	0.76 (4.5)	0.40 (5)	0.00 (4.5)	0.20 (4)	0.88 (5)	0.92 (5)	0.90 (4.5)	0.91 (2)
	IG	0.99 (1.5)	0.99 (2)	0.99 (1.5)	0.87 (5)	0.99 (2.5)	1.00 (1)	0.90 (2.5)	0.98 (3)	0.76 (2.5)	0.97 (1)	1.00 (1)	0.44 (3)	0.99 (1.5)	0.99 (1)	0.99 (1.5)	0.86 (4.5)
	ttest	0.97 (3)	0.98 (3)	0.91 (3)	0.88 (4)	0.99 (2.5)	0.99 (2)	0.89 (5)	1.00 (1.5)	0.77 (4.5)	0.80 (3)	0.15 (3)	0.52 (2)	0.97 (3)	0.98 (2.5)	0.91 (3)	0.86 (4.5)
	P20	0.99 (1.5)	1.00 (1)	0.99 (1.5)	0.93 (1)	1.00 (1)	0.98 (3)	0.90 (2.5)	1.00 (1)	0.95 (1)	0.90 (2)	0.90 (2)	0.65 (1)	0.99 (1.5)	0.98 (2.5)	0.99 (1.5)	0.93 (1)
DLBCL stanford	SNR	0.70 (4)	0.69 (5)	0.63 (5)	0.69 (5)	0.88 (4.5)	0.77 (5)	0.50 (5)	0.79 (5)	0.66 (4)	0.18 (5)	0.22 (5)	0.35 (5)	0.82 (3)	0.71 (4.5)	0.28 (5)	0.70 (4.5)
	PCC	0.71 (2)	0.71 (3)	0.70 (3.5)	0.70 (3.5)	0.96 (2.5)	0.93 (4)	1.00 (2.5)	0.85 (3.5)	0.68 (2)	0.65 (2.5)	0.66 (3)	0.57 (3.5)	0.85 (2)	0.71 (4.5)	0.79 (3)	0.70 (4.5)
	IG	0.70 (4)	0.71 (3)	0.71 (2)	0.70 (3.5)	0.88 (4.5)	0.97 (3)	1.00 (2.5)	0.85 (3.5)	0.52 (5)	0.65 (2.5)	0.69 (2)	0.57 (3.5)	0.75 (5)	0.86 (1)	0.85 (2)	0.77 (3)
	ttest	0.70 (4)	0.71 (3)	0.70 (3.5)	0.71 (2)	0.96 (2.5)	0.96 (1.5)	1.00 (2.5)	0.98 (2)	0.67 (3)	0.65 (2.5)	0.46 (4)	0.66 (2)	0.81 (4)	0.77 (3)	0.75 (4)	0.83 (1)
	P20	0.94 (1)	0.85 (1)	0.90 (1)	0.83 (1)	0.99 (1)	0.96 (1.5)	1.00 (2.5)	1.00 (1)	0.88 (1)	0.65 (2.5)	0.82 (1)	0.67 (1)	0.92 (1)	0.81 (2)	0.87 (1)	0.78 (2)

Table 17
Average grade of 1st order grading of P20 and all single ranking method.

RankingModel	Accuracy				Sensitivity				Specificity				Jaccard			
	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR	KNN	NB	ANN	MLR
SNR	3.57	4.29	4.21	3.93	4.14	3.86	4.14	4.29	3.86	4.07	4.64	4.36	3.71	3.64	4.43	4.21
PCC	2.86	3.14	3.57	3.00	3.07	3.00	3.14	3.86	3.50	3.86	3.50	3.14	2.86	3.57	3.50	3.00
IG	3.07	3.00	3.07	3.57	2.71	2.29	2.64	2.50	2.71	2.14	2.00	3.29	3.29	2.00	2.21	3.36
ttest	3.86	3.43	3.07	3.21	3.43	2.86	2.86	2.29	3.00	2.50	2.79	2.21	3.14	3.07	2.86	2.50
P20	1.50	1.14	1.07	1.29	1.50	2.71	2.20	2.07	1.93	2.43	2.07	2.00	2.00	2.57	2.00	1.93

Table 18
Second order grading of P20 and all single ranking method in different performance metrics.

RankingModel	Accuracy					Sensitivity					Specificity					Jaccard				
	KNN	NB	ANN	MLR	Average Grade	KNN	NB	ANN	MLR	Average Grade	KNN	NB	ANN	MLR	Average Grade	KNN	NB	ANN	MLR	Average Grade
SNR	4.00	5.00	5.00	5.00	4.75	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
PCC	2.00	3.00	4.00	2.00	2.75	3.00	4.00	4.00	4.00	3.75	4.00	4.00	4.00	3.00	3.75	2.00	4.00	4.00	3.00	3.25
IG	3.00	2.00	2.50	4.00	2.88	2.00	1.00	2.00	3.00	2.00	2.00	1.00	1.00	4.00	2.00	4.00	1.00	2.00	4.00	2.75
ttest	5.00	4.00	2.50	3.00	3.63	4.00	3.00	3.00	2.00	3.00	3.00	3.00	2.00	2.75	3.00	3.00	3.00	2.00	2.75	
P20	1.00	1.00	1.00	1.00	1.00	1.00	2.00	1.00	1.00	1.25	1.00	2.00	2.00	1.00	1.50	1.00	2.00	1.00	1.00	1.25

pared. Since those performance measurements are not guaranteed to be normally distributed, Nemenyi post-hoc test is suggested [33]. Nemenyi post hoc test is a nonparametric test to find out the significance difference among the model. It works on the null hypothesis that there exist no significance difference between a pair of models and the null hypothesis is rejected if significance difference exists. In this test pair wise comparison of model is done to find whether a significant difference exists between two models. If grade of the models differ by the critical difference (CD), then a significant difference exists. The critical difference is defined as

$$CD = q_{\alpha} \sqrt{\frac{k(k+1)}{6N}} \quad (12)$$

Where the value of q_{α} is the critical value of Q for a multiple non-parametric comparison with a control (Table B.16 in [34]), k is number of models and N is number of classifier.

Considering $\alpha = 0.10$, q_{α} is found to be 3.417 and the corresponding CD is $3.417 * \sqrt{\frac{24*25}{6*4}} = 17.08$. Nemenyi post hoc test is used for pair wise comparison. If the grade difference between the worst performing model and model under consideration is larger than the CD value then the post-hoc test is powerful enough to detect any significant differences between the models. In this case study, the significance difference is obtained from the difference of average grade of a model and the average grade of worst model. Considering each performance metric significance difference is evaluated and shown for three pipeline models in Table 15.

The "*" marked values specified in Table 15 are less than the CD value, which indicates significant difference does not exist. It is observed that for sensitivity performance metric of P18 and for specificity and jaccard performance metrics of P2 significant difference does not exist. Only P20 pipeline model satisfies the Nemenyi post hoc test for significant difference for all performance metrics. Apart from these three pipelines models presented in Table 15, all other models do not satisfy Nemenyi post hoc test for all or majority of the performance metrics. Therefore model P20 can be treated as a consistent model, P20 may not perform the best always but its results will be competitive with other better performing models.

Further the performance metrics of model P20 and ranking techniques along with their respective grades are presented in Table 16.

In this table, the highest performance metric values for each performance metric/classifier/database are made bold faces. However considering seven databases, four classifiers and four performance metrics, there are 112 tests for P20; out of which in 70 tests i.e.

in 63% tests it yield highest performance. Further out of these 70 highest results, in 27 tests few other techniques have also yield the same result as P20, whereas in 43 test cases P20 has found highest performance, which is not achieved by any single ranking technique. Still then complete mandate is not in favour of P20, therefore statistical significance test of these results is essential to ensure that P20 is significantly better technique than others or not. Further as the performance of these ranking technique depends on both classifier as well as on the dataset, single stage grading is not good enough to test the statistical significance of the results of four ranking techniques and model P20. Therefore 1st stage grading of the ranking techniques is shown in Table 17 along with the results for each dataset/classifier. As discussed before, the highest performance of a ranking technique for a database/classifier is assigned grade 1, next highest as 2 and so on. When more than one ranking technique has the same level of performance, average grade is assigned to each one. For example in the third column of Table 16 for Leukemia database, PCC and IG yield the highest accuracy of 0.98 for KNN, their average grade $(1+2)/2 = 1.5$ is assigned to both. Next highest accuracy of 0.95 is yield by P20 is assigned grade 3, next highest accuracy of 0.94 is yield by test is assigned 4 and the minimum accuracy is yield by SNR is assigned grade 5.

Table 17 shows the average grade of the databases of the 1st order grading of Table 16. For example if the average grade of SNR for accuracy/KNN in Table 18 is referred as $T_{17}(\text{SNR}, \text{Accuracy}, \text{KNN})$. Then

$$T_{17}(\text{SNR}, \text{Accuracy}, \text{KNN}) = [T_{16}(D_1, \text{SNR}, \text{Accuracy}, \text{KNN}) + T_{16}(D_2, \text{SNR}, \text{Accuracy}, \text{KNN}) + \dots + T_{16}(D_7, \text{SNR}, \text{Accuracy}, \text{KNN})]/7,$$

where T_{16} refers to Table 16 and D_1, D_2, \dots, D_7 refers to Leukemia, Colon, CNS, Prostrate, Lungcancer1, Lungcancer2, and DLBCL Stanford databases, respectively.

From the average grades of the 1st stage grading in Table 17, the 2nd stage grading of the ranking techniques is obtained and the 2nd stage grades are presented in Table 18. For example taking average grade of accuracy for KNN classifier in Table 17, the lowest average grade is 1.5 of pipeline P20, which is assigned grade 1. The next lowest average grade is 2.86 of PCC, which is assigned grade 2, and so on.

In Table 18, the last column under each performance metric shows the average grades of the four classifiers for the ranking model. To validate the results, Nemenyi post hoc test is applied for pair wise comparison, where q_{α} is found to be 2.29 and the

Table 19
Comparison of computation time (in seconds) of single ranking system with P20.

Databases	KNN					NB					ANN					MLR				
	SNR	PCC	IG	ttest	P20	SNR	PCC	IG	ttest	P20	SNR	PCC	IG	ttest	P20	SNR	PCC	IG	ttest	P20
CNS	12.6	0.7	14	1.14	28.4	8.5	0.7	4.6	1.17	27.8	32.3	15	60	15	41	8.9	0.66	15	1.2	27.6
Leukemia	5.03	0.5	16	0.91	17.6	4.9	0.5	7.2	0.92	17.4	13.4	8.9	8.8	9.3	26	4.9	0.52	13	0.9	17.5
Postrate	24.1	1.7	18	2.88	48.4	26	1.6	19	2.86	66.4	35.4	14	19	16	80	35	1.7	18	2.9	56.8
DLBCLStanford	16.4	0.9	27	1.43	41.8	11	0.9	25	1.4	40.4	38.2	17	26	18	56	11	0.91	24	1.4	41.7
Lungcancer1	62.1	4.1	62	7.09	150	62	3.9	61	11.9	151	98.1	37	65	47	186	61	3.8	62	94	151
Lungcancer2	22.6	1.6	13	2.27	49.8	13	1.2	10	1.72	52.5	34.8	22	14	24	71	15	1.1	19	1.8	24.8
Colon Cancer	5.85	0.2	3.6	0.21	7.08	4	0.2	3.9	0.19	7.07	3	13	4.1	14	21	3.6	0.15	4	0.2	7.06

corresponding CD is 2.09. The significance difference between the worst ranking technique and P20 is evaluated for each performance metric. The significance difference obtained as follows.

- i Significance difference of P20 for accuracy: $4.75 - 1.00 = 3.75$
- ii Significance difference of P20 for sensitivity: $5.00 - 1.25 = 3.75$
- iii Significance difference of P20 for specificity: $5.00 - 1.50 = 3.50$
- iv Significance difference of P20 for Jaccard: $5.00 - 1.25 = 3.75$

In case of all the performance metrics the significant difference of P20 is higher than the CD value of 2.09; therefore it satisfies the Nemenyi post hoc test. No other ranking technique satisfies the Nemenyi post hoc test for all the performance metrics. Therefore P20 can be considered as a better model in comparison to four ranking techniques.

Further the time for computation required by the pipeline model P20 is compared with that of single ranking techniques and the results are presented in Table 19. As expected the model P20 requires much more time than that of any of the single ranking technique and in few cases it requires time more than the sum of time required by all the single ranking techniques. In single ranking technique, the database with large feature space is used only once to evaluate and rank the features, and then hardly 5–10 features are selected for the next steps of action. Whereas in pipelining technique, after ranking in each stage, new database is generated with 80% of the features, which is also quite large and requires lot of time to manipulate.

However, the results of this model intended to contribute to the detection of disease of a human being or related issues, therefore the additional time required to have a better prediction is not beyond tolerance.

5. Conclusion & future work

The contribution of proposed work is outlined as follows:

- In this case study, a pipeline of gene ranking method is applied to eliminate some less significant genes at each stage of the pipeline.
- Each of the ranking techniques used for the pipeline possesses different evaluation criterion and ranking system. Therefore very often a particular gene is assigned different rank by different ranking technique, which helps different pipelines to eliminate different set of genes at different stages and the set of genes selected after passing through different pipelines are always different.
- For this work four gene ranking method with significant difference in ranking approach considered such as signal to noise ratio, pearson correlation coefficient, information gain and t-statistic to construct the pipeline.
- The combination of four ranking method generates twenty four unique pipeline models, performance of which is evaluated and compared for seven publicly available gene expression databases.

- Further to overrule the performance of single classifier leading to biasness, four different classifiers are used such as MLR, ANN, naïve Bayesian network and kNN.
- Again the loophole of a single performance metric is taken care by considered four performance metrics for evaluation of these models.
- From the simulation results it is revealed that finding a single pipeline model or group of models, suitable for best performance is difficult because performance of no single pipeline or a group dominates the rest. Therefore the technique of grading the pipeline models is employed to find out a competitive pipeline model.
- To validate the effectiveness of the proposed method Nemenyi post hoc nonparametric statistical test is conducted to determine the significant difference among these models. Finally it is observed that the sequence of feature ranking techniques of the pipeline model P20 i.e. [correlation coefficient → SNR → t-statistic → information gain] is found to be the most effective in comparison to all other pipelines.
- The performance of pipeline model P20 is also compared with the performance of each single ranking technique considered for construction of this pipeline. It is observed that P20 does not yield the best result always in comparison to single ranking techniques, but satisfies the statistical test of significance among the single ranking techniques. Further in 63% instances P20 has obtained highest level of performance and in 38% cases it dominates the performance of all other single ranking techniques.
- In majority of cases the computation time required by model P20 is much higher in comparison to all other single ranking techniques. However the application of such model is not intended for any real time system, therefore the additional computation time required can be accommodated for using a consistent model.

This approach improves the efficiency of the classifier with minimal computational effort. So future research work can be extended to implement other feature selection technique like filter or wrapper based approach. Performance improvement of proposed model can be done using other classifier and performance metrics. In future this model can also be applied for other high dimensional data for further validation.

References

- [1] H.W. Resson, R.S. Varghese, Z. Zhang, J. Xuan, R. Clarke, Classification algorithms for phenotype prediction in genomics and proteomics, *Front. Biosci.* 13 (2008) 691–708.
- [2] Kun-Hong Liu, Bo Li, Qing-Qiang Wu, Jun Zhang, Ji-Xiang Du, Guo-Yan Liu, Microarray data classification based on ensemble independent component selection, *Comput. Biol. Med.* 39 (11) (2009) 953–960.
- [3] M. Dash, H. Liu, Feature selection for classification, *Intell. Data Anal.* 1 (3) (1997).
- [4] Hui Wang, David Bell, Fionn Murtagh, Relevance approach to feature subset selection, in: Feature Extraction, Construction and Selection, Springer, US, 1998, pp. 85–99.

- [5] R. Kohavi, G.H. John, Wrapper for feature subset selection, *Artif. Intell.* 97 (12) (1997) 273–324.
- [6] K. Kira, L.A. Rendell, A practical approach to feature selection, in: *Proceedings of the Ninth International Workshop on Machine Learning*, Morgan Kaufmann Publishers Inc., Scotland, United Kingdom, 1992, pp. 249–256.
- [7] Stuart W. Card, Information distance based fitness and diversity metrics, in: *Proceedings of the 12th Annual Conference Companion on Genetic and Evolutionary Computation*, ACM, 2010, pp. 1851–1854.
- [8] Boyan Bonev, Francisco Escolano, Daniela Giorgi, Silvia Biasotti, Information-theoretic selection of high-dimensional spectral features for structural recognition, *Comput. Vision Image Understand.* 117 (3) (2013) 214–228.
- [9] Yvan Saeys, Iñaki Inza, Pedro Larrañaga, A review of feature selection techniques in bioinformatics, *Bioinformatics* 23 (19) (2007) 2507–2517.
- [10] V. Bolón-Canedo, N. Sánchez-Maróño, A. Alonso-Betanzos, J.M. Benítez, F. Herrera, A review of microarray datasets and applied feature selection methods, *Inf. Sci.* 282 (2014) 111–135.
- [11] Y. Saeys, I. Inza, P. Larrañaga, A review of feature selection techniques in bioinformatics, *Bioinformatics* 23 (19) (2007) 2507–2517.
- [12] Lazar Cosmin, Jonatan Taminau, Stijn Meganck, David Steenhoff, Alain Coletta, Colin Molter, Virginie de Schaetzen, Robin Duque, Hugues Bersini, Ann Nowe, A survey on filter techniques for feature selection in gene expression microarray analysis, *IEEE/ACM Trans. Comput. Biol. Bioinform.* 9 (4) (2012) 1106–1119.
- [13] I. Guyon, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
- [14] G.V.S. George, V.C. Raj, Review on Feature Selection Techniques and the impact of SVM for cancer Classification using Gene expression Profile, *Int. J. Comp. Sci. Eng. Surv.* 2 (2011) 16–27.
- [15] G. Chandrashekar, F. Sahin, A survey on feature selection methods, *Comput. Electr. Eng.* 40 (2014) 16–28.
- [16] Mark A. Hall, Lloyd A. Smith, Practical feature subset selection for machine learning (1998).
- [17] Peyman Jafari, Francisco Azuaje, An assessment of recently published gene expression data analyses: reporting experimental design and statistical factors, *BMC Med. Inform. Decis. Mak.* 6 (1) (2006) 27.
- [18] Jing-jing Zhao, Xiao-hong Huang, S.U.N. Qiong, M.A. Yan, Real-time feature selection in traffic classification, *J. China Univ. Posts Telecommun.* 15 (2008) 68–72.
- [19] Jeffrey G. Thomas, James M. Olson, Stephen J. Tapscott, Lue Ping Zhao, An efficient and robust statistical modeling approach to discover differentially expressed genes using genomic expression profiles, *Genome Res.* 11 (7) (2001) 1227–1236.
- [20] Chen A. Tsai, Yi Ju Chen, James J. Chen, Testing for differentially expressed genes with microarray data, *Nucleic Acids Res.* 31 (9) (2003) e52.
- [21] A. Antoniadis, S. Lambert-Lacroix, F. Leblanc, Effective dimension reduction methods for tumor classification using gene expression data, *Bioinformatics* 19 (2003) 563–570.
- [22] D. Hwang, W.A. Schmitt, G. Stephanopoulos, G. Stephanopoulos, Determination of minimum sample size and discriminatory expression patterns in microarray data, *Bioinformatics* 18 (2002) 1184–1193.
- [23] Z. Wang, *Neuro-Fuzzy Modeling for Microarray Cancer Gene Expression Data*, University Computing Laboratory, Oxford, 2005.
- [24] T.R. Golub, D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, H. Coller, M.L. Loh, J.R. Downing, M.A. Caligiuri, C.D. Bloomfield, E.S. Lander, Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, *Science* 286 (1999) 531–537.
- [25] Sung-Bae Cho, Hong-Hee Won, Machine learning in DNA microarray analysis for cancer classification, *Proc. First Asia Pac. Bioinform. Conf. Bioinform.* 2003 (19) (2003) 189–198.
- [26] H. Hu, J. Li, H. Wang, G. Daggard, Combined gene selection methods for microarray data analysis, in: *Knowledge-Based Intelligent Information and Engineering Systems*, Springer, Berlin Heidelberg, 2006, pp. 976–983.
- [27] Eric P. Xing, Michael I. Jordan, Richard M. Karp, Feature selection for high-dimensional genomic microarray data, *ICML* 1 (2001) 601–608.
- [28] Y. Peng, W. Li, Y. Liu, A hybrid approach for biomarker discovery from microarray gene expression data for cancer classification, *Cancer Inf. J.* 2 (2007) 301–311.
- [29] P.A. Mundra, J.C. Rajapakse, Gene and sample selection for cancer classification with support vectors based t-statistic, *Neurocomputing* 7 (13) (2010) 2353–2362.
- [30] C. Lee, Y. Leu, A novel hybrid feature selection method for microarray data analysis, *Appl. Softw. Comput. J.* 11 (1) (2011) 208–213.
- [31] T. Abeel, T. Helleputte, Y. Van de Peer, P. Dupont, Y. Saeys, Robust biomarker identification for cancer diagnosis with ensemble feature selection methods, *Bioinformatics* 26 (3) (2000) 392–398.
- [32] F. Yang, K. Mao, Robust feature selection for microarray data based on multicriterion fusion, *IEEE/ACM Trans. Comput. Biol. Bioinform.* 8 (4) (2011) 1080–1092.
- [33] J. Demsar, Statistical comparisons of classifiers over multiple data sets, *J. Mach. Learn. Res.* 7 (2006) 1–30.
- [34] More on dichotomous variables, in: J.H. Zar (Ed.), *Biostatistical Analysis*, 4th ed., Prentice Hall, Upper Saddle River, 1999, pp. 516–565.
- [35] T.R. Golub, Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, *Science* 286 (1999) 531–537.
- [36] David G. Beer, Gene-expression profiles predict survival of patients with lung adenocarcinoma, *Nat. Med.* 8 (8) (2002) 816–823.
- [37] Gavin J.Gordon, Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma, *Cancer Res.* 62 (2002) 4963–4967.
- [38] Scott L. Pomeroy, Prediction of central nervous system embryonal tumour outcome based on gene expression, *Nature* 415 (2002) 436–442.
- [39] Dinesh Singh, Gene expression correlates of clinical prostate cancer behavior, *Cancer Cell* 1 (2002) 203–209.
- [40] Ash A. Alizadeh, Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling, *Nature* 403 (2000) 503–511.
- [41] Rasmita Dash, et al., Efficient Microarray Data Classification with Three-Stage Dimensionality Reduction Intelligent Computing, Communication and Devices, Springer, India, 2015, pp. 805–812.
- [42] Rasmita Dash, B.B. Misra, Reduction Combination Determination for Efficient Microarray Data Classification with Three Stage Dimensionality Reduction Approach. *Computational Intelligence in Data Mining-Volume 2*, Springer, India, 2015, pp. 695–703.
- [43] M. Yassi, H.M. Mohammad, Robust and stable feature selection by integrating ranking methods and wrapper technique in genetic data classification, *Biochem. Biophys. Res. Commun.* 446 (2014) 850–856.
- [44] S. Cho, H. Won, Data mining for gene expression profiles from DNA microarray, *Int. J. Softw. Eng. Knowl. Eng.* 13 (6) (2003) 593–608.
- [45] Jie Xiang, et al., A novel hybrid system for feature selection based on an improved gravitational search algorithm and k-NN method, *Appl. Soft Comput.* 31 (2015) 293–307.
- [46] V. Elyasigomari, M.S. Mirjafari, H.R.C. Screen, M.H. Shaheed, Cancer classification using a novel gene selection approach by means of shuffling based on data clustering with optimization, *Appl. Soft Comput.* 35 (2015) 43–51.
- [47] Sudipto Chaki, B. Shanmugarajan, Sujit Ghosal, G. Padmanabham, Application of integrated soft computing techniques for optimisation of hybrid CO₂ laser-MIG welding process, *Appl. Soft Comput.* 30 (2015) 365–374.
- [48] Sudipto Chaki, Ravi N. Bathe, Sujit Ghosal, G. Padmanabham, Multi-objective optimisation of pulsed Nd: YAG laser cutting process using integrated ANN-NSGAI model, *J. Intell. Manuf.* (2015) 1–16.
- [49] S. Chaki, S. Ghosal, A. Ga-Ann, Hybrid model for prediction and optimization of Co₂ laser-Mig hybrid welding process, *Int. J. Automot. Mech. Eng.* 11 (2015).
- [50] V. Bolón-Canedo, N. Sánchez-Maróño, A. Alonso-Betanzos, Distributed feature selection: an application to microarray data classification, *Appl. Soft Comput.* 30 (2015) 136–150.
- [51] Su Sin Chong, A.R. Abdul Aziz, Sulaiman W. Harun, Hamzah Arof, Shahabuddin Shamshirband, Application of multiple linear regression, central composite design, and ANFIS models in dye concentration measurement and prediction using plastic optical fiber sensor, *Measurement* 74 (2015) 78–86.
- [52] S.C. Satapathy, J.V.R. Murty, P. Reddy, B.B. Misra, P.K. Dash, G. Panda, Particle swarm optimized multiple regression linear model for data classification, *Appl. Soft Comput.* 9 (March (2)) (2009) 470–476.