Australian Journal of Basic and Applied Sciences, 9(27) August 2015, Pages: 145-151

ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

A Comparative Study of different Classifiers on Microarray Cancer Gene Expression Data

[1]C.Arunkumar and [2]Dr.S.Ramakrishnan

[1]*Assistant Professor(Senior Grade), Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, Tamilnadu, India*
[2]*Professor and Head, Department of Information Technology, Dr.Mahalingam College of Engineering and Technology, Pollachi, Tamilnadu, India*

ARTICLE INFO

Article history:
Received 23 June 2015
Accepted 25 August 2015
Available online 2 September 2015

Keywords:
Naïve Bayes, Random Forest, Bagging, Functional Tree, Lib Linear, LibSVM

ABSTRACT

Microarray gene expression data has numerous clinical applications and cancer diagnosis and classification is one of the emerging applications. However, the main challenge in the classification of the disease is the availability of fewer training and testing samples and a very large number of gene information. In this paper, a survey of the performance of six classification algorithms is analyzed using three different datasets after the process of feature selection is performed using a hybrid approach that combines correlation based feature selection with linear forward selection. We focus our attention on 3 different cancer datasets that uses a binary classifier to predict the accuracy of the system. For the purpose of classification, we focus on Bayes Classifier, functional classifier, Meta classifier and Tree based classifier. The classification accuracy of the different classifiers on the cancer datasets are analyzed and presented.

© 2015 AENSI Publisher All rights reserved.

To Cite This Article: C.Arunkumar and Dr.S.Ramakrishnan., A Comparitive Study of different classifiers on microarray cancer gene expression data. ***Aust. J. Basic & Appl. Sci.,*** *9(27): 145-151, 2015*

INTRODUCTION

The predictive models have been widely used by statisticians to infer the dependencies of the variables on a set of observed data. Microarray analysis involves the use of a very large dataset with thousand of gene information, large amount of noise, non linear dependencies on the observed data and very few samples for analysis purposes. One of the main applications of microarray analysis is cancer classification in the medical field. This involves analyzing the gene information and classifying the sample as normal or diseased tissue. The key challenge in microarray analysis is the availability of a large number of features usually in the order of thousands and very few samples in the order of hundreds. The functional relationships between the genes need to be identified for designing a suitable classifier. The numerical accuracy of the classifier could be improved only by reducing the number of features. Selection of an appropriate subset of genes is essential for effective patient management and therapy (M. J. van de Vijver *et al.*,2002; L. J. van 't Veer *et al.*,2002).

The process of gene selection is called Feature Selection in machine learning. The main objective of feature selection is to select a set of genes that would probably result in a better predictive model. The intelligibility of the model is increased and the storage requirements and measurements are decreased as a result of feature selection (I. Guyon and A. Elisseeff, 2003). Redundant and irrelevant features greatly affect the predictive accuracy of the model(A. Blum and P. Langley, 1997)(R. Kohavi and G. H. John, 1997)(G. Provan and M. Singh, 1995). This fact is clearly evident from the experimental studies carried out on our dataset. The filter and wrapper approaches are the two well-known methods in feature selection (A. Blum and P. Langley, 1997)(R. Kohavi and G. H. John, 1997). In the filter approach, some examples of evaluation functions are probabilistic distance, interclass distance, information-theoretic or probabilistic dependence measures(M. Dash and H. Liu, 2000). The above measures called as the intrinsic properties, are calculated from raw data instead of employing a learning model that smoothes or reduce the noise. The validation outcome of the learning algorithm is an evaluation function in the wrapper approach.

Corresponding Author: C.Arunkumar, Assistant Professor(Senior Grade), Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, Tamilnadu, India
E-mail: c_arunkumar@cb.amrita.edu

The filter and wrapper approaches exhibit lot of disadvantages. To overcome the problems of the filter and wrapper approaches, hybrid filter-wrapper models have been proposed. The simplicity nature of the filter model in gene selection is combined with the wrapper model that could optimize the classification accuracy in final feature selection. The hybrid model works in two stages. In the first stage, the filter model is applied to remove the irrelevant genes from the dataset and hence the size of the dataset becomes manageable for the second stage(Yukyee Leung and Yeungsam Hung, 2010). In the second stage, the wrapper model is applied on the reduced feature subset that optimizes the accuracy in training. The embedded approach seeks to subsume feature selection as part of the model building process. Therefore, this approach is associated with a specific learning algorithm. On the other hand, the goal of the hybrid approach is to take advantage of both the filter and the wrapper approaches (Sina Tabakhi *et al*., 2014).

MATERIALS AND METHODS

Dataset Used:

The dataset used for this survey is downloaded from the Kentridge Biomedical Repository. Three different cancer datasets namely ALL/AML, Ovarian Cancer and Lung Cancer are used for this study (http://datam.i2r.a-star.edu.sg/datasets/krbd/). The table below shows the total number of genes, the number of genes used in the training and testing dataset. These datasets consists of binary class namely healthy and diseased tissues.

Table 1: Description of the Dataset used for the study

Name of the Disease	Total Number of Genes	Number of Training Samples	Number of Testing Samples
ALL/AML	7,129	43	29
Ovarian Cancer	15,154	152	101
Lung Cancer	12,533	109	73

Classifiers Used:

The datasets obtained from the Kentridge Biomedical Repository are analyzed and classified using the WEKA(http://www.cs.waikato.ac.nz/ml/weka/), the collection of machine learning algorithms for data mining tasks. There are 6 different classifiers that are used to analyze and classify the microarray gene expression data. The different classifiers fall under the category of Bayes, Functions, Meta and Trees. The diagrammatic representation of the different classifiers is given in Fig.1 below:

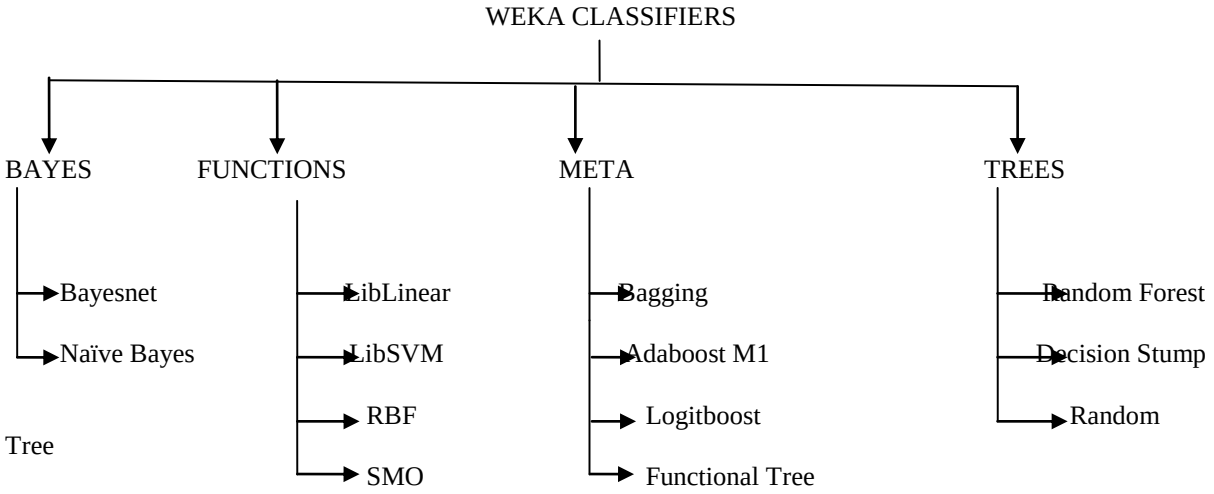

Fig. 1: WEKA Classifiers

Methods:

Naive Bayes Classifier:

A naïve Bayes classifier is a simple probabilistic classifier that follows the independent feature model. This model uses the Bayes theorem as a basis that employs strong independent assumptions. It is also one of the oldest classifiers that uses the Bayes rule for classification(Pradipta Maji, 2012).

In simple terms, a naive Bayes classifier assumes that the presence of a particular feature of a class is unrelated to the presence of any other feature and vice-versa. For example, if the classification problem involves classifying the different types of fruits, an apple could be identified based on certain features like its shape (round), colour (red) etc.,(Dr.A.Kumaravel and A.Aarthi, 2013). Even if these features are dependent among themselves or on other features, this classifier considers all of these properties to independently contribute to the fact that the fruit is an apple. An important part of this algorithm is that it is independent of the features and

produces excellent results (Qinbao Song *et al*., 2013).

In a supervised learning environment, the naïve bayes classifier could be trained very efficiently. This requires knowledge of the nature of the probability model. Many practical applications involve the usage of the Maximum likelihood method. In such cases, the model could work perfectly without using any of the Bayesian methods or the concept of Bayesian probability.

Naïve bayes classifiers are most suitable for complex real world problems. This has been true even though the classifier has a naïve design and all assumptions made are over simplified. Even though the analysis of the Bayesian classification problems in 2004 showed some unreasonable efficacy, this classifier outperforms more current approaches like trees and random forests (http://www.princeton.edu/~achaney/tmve/wiki100k/docs/Naive_Bayes_classifier.html).

The major advantage of the Naïve Bayes Classifier is that it requires fewer amounts of data during the training phase. This small amount of data is sufficient to estimate the mean and variance during the process of classification. It does not require the calculation of the entire covariance matrix. Only the variances of the variables of each class needs to be computed because of the assumption of the use of independent variables (Shadab Adam Pattekari and Asma Parveen, 2012).

Random Forest:

Many classification trees are grown by random forests. A new object is classified from the set of input vectors by putting the input vector down each of the trees in the forest. The process of 'voting' is carried out for the purpose of classification. The forest chooses the classification having the most votes over all the trees in the forest (Florin Leon *et al*., 2010).

The strength of individual trees in a forest and Correlation between any two trees are two main factors that contribute to the forest error rate. The forest error rate is directly proportional to the correlation value and inversely proportional to the strength of the individual tree. A lower error rate is desired to obtain a strong classifier. Random forest algorithm is highly sensitive to the 'm' parameter, because when the values of 'm changes, the correlation and strength also changes. 'm' is the only adjustable parameter in the classifier. Hence it becomes essential to find an optimal value for 'm'. The Out-of-Bag(OOB) error rate aids in finding an optimal value for 'm. Some of the features of Random Forest include highest accuracy among the available algorithms, efficient execution in large datasets, ability to handle thousands of variables without variable deletion, ability to highlight the important variables in a classification, generation of an internal unbiased estimate of the generalization error as the forest building progresses, error balancing in unbalanced datasets, implementation of effective methods to populate the missing data and maintaining the accuracy at the same time, detecting variable interactions, computation of proximity between pairs of cases that can be used in clustering, computation of prototypes that provide information about the relation between variables, the ability to apply this technique to label the unsupervised data and to detect the presence of outliers. Two data objects are generated by random forests. Sampling with replacement is performed on the training set for the current tree. In that case, about one-third of the cases will be left out of the sample. As trees are added to the forest, this Out-of-Bag data (OOB) would be useful to obtain an unbiased estimate in the classification error (Jian-Hua Huang *et al*., 2013). Some of the issues with high dimensional data are missing values, outliers and low dimensional views. These issues could be resolved by the usage of proximities. During the process of tree building, the entire set of data is made to run down the tree and the corresponding proximity values for each pair are computed. The values of proximity are incremented by 1 if two pairs occupy the same terminal node (http://www.stat.berkeley.edu/users/breiman/RandomForests/cc_home.htm).

Bagging:

Bagging or bootstrap aggregation is an ensemble method that creates individuals for its ensemble from the training dataset. The process of random redistribution is carried out during training. Assume the size of the original training set to be 'N'. Bagging generates the training set by drawing 'M' samples randomly from 'N' with replacement. This causes some samples to be repeated again while some may be left out during training. Different random sampling of the training set is used to generate individual classifiers in the ensemble (Gulshan Kumar and Krishan Kumar, 2012)(M. Govindarajan and RM. Chandrasekaran, 2012). The Bagging method is found to be very effective on unstable learning algorithms. In this case, a small change in the training set results in large changes in the predictions. Neural Networks and Decision trees are examples of unstable learning algorithms. Bootstrap is completely automatic process. No theoretical calculations are required for the Bootstrap process. It is not based on asymptotic results.

The prediction accuracy could be improved by using the bootstrap that takes on three steps to achieve its goal namely creation of the sample, estimation of the underlying model for each of the bootstrap samples and aggregation of the predictions.

Functional Tree:

An intuitive appeal is created whenever an algorithm is designed that explores multiple representation languages and different search spaces

(Harry Hochheiser and Ben Shneiderman, 2001). A combination of different attributes is used to generate decision test results that are used to explore multiple representation languages. This is applicable in the context of classification problems. Model-tree algorithms in regression domain use a similar approach, but they make use of linear models at leaf nodes (Joao Gama, 2004). This algorithm makes use of the decision nodes that employ multivariate tests. The discriminant function is used to predict the class of the leaf nodes. The two common operations employed on trees are growth and pruning. Functional leaves are an outcome of pruning the tree and Multivariate decision nodes are built when the tree is grown. Multivariate trees are a specialized class of functional trees (Gama, J, 2001).

Lib Linear:

Linear Classification has been found to be one of the promising techniques for classification of large sparse data that contains huge number of features and their instances. The LibLinear package consists of a library of functions and command line tools that would perform the process of learning and classification. The Lib Linear package is an open source tool for large scale data classification. This package resembles the LibSVM in terms of its API but differs in terms of its model after the process of training is performed on the dataset (Rong-En Fan *et al*., 2008). The main features of LibLinear are it could be used for multi-class classification, cross validation could be performed on the data, logistic regression could be estimated, weights could be calculated for unbalanced data and it supports multiple interfaces in MATLAB, Ruby, Python etc., It also taken on the same data format as LibSVM package and hence it is easy to learn (http://www.csie.ntu.edu.tw/~cjlin/liblinear/).

Lib SVM:

Support vector classification, distribution estimation and regression can be performed by an integrated package called Lib SVM. It supports multi-class classification (http://www.csie.ntu.edu.tw/~cjlin/libsvm/). LibSVM has certain features that include probability estimates, wide variety of kernels, different SVM formulations cross validation for model selection and weighted SVM for unbalanced data.

Correlation-based Feature Selection:

The Correlation-based Feature Selection (CBFS) is a multivariate ranking technique that identifies the correlation that exists between different gene subsets in a dataset. It differs from the univariate technique in the sense that the univariate technique does not take into account the interaction between the features (Mark.A.Hall, 1999)(Cosmin Lazar *et al*., 2012)(Li M. Fu and Eun Seog Youn, 2003). The attribute subsets are ranked based on a correlation based heuristic evaluation function. This heuristic evaluation function is then used to compute the correlation coefficient. A subset of attributes is evaluated by considering the identification ability of each attribute. The identification ability of each of the attributes is used to evaluate a subset of attributes. Pearsons correlation coefficient is very sensitive to the presence of outliers and noise (Joaquim F. Pinto da Costa *et al*., 2011). The relationship between variables (Genes) can be measured by the process of correlation (Cheng-San Yang *et al*., 2008). Pearson Product Moment Correlation (PPMC) is the most common measure of correlation in statistics, which shows the linear relationship between two variables. Formula for calculating Pearson correlation between features x_i and y_i is given in Eq 1

$$Correlation = \left(\Sigma\ (\mathrm{xi} - \mathrm{mean}\ (\mathrm{xi}) * \frac{\mathrm{yi} - \mathrm{mean}\,(\mathrm{yi})}{\mathrm{n} * \mathrm{Standard\ deviation}(\mathrm{xi}) * \mathrm{Standard\ deviation}\,(\mathrm{yi})}\right) \quad \text{Eq 1}$$

Pearson correlation coefficient between attributes is found out. Attributes having low inter-correlation are selected (Ammu Prasanna Kumar and Preeja Valsala, 2013). The WEKA tool is used to implement CBFS. The selected genes were used to study the different types of cancer. The attributes exhibit high correlation if the value of correlation coefficient lies between 0.5 and 1 and is said to be less correlated if its value lies between 0.3 and 0.5.

Best first, forward selection and backward elimination are some of the common methods used in Correlation Based Feature Selection. In this study, we used Weka to implement CFS, and used selected gene subsets to identify different kinds of cancer types and various diseases (Cheng-San Yang *et al*., 2008)(Li-Yeh Chuang *et al*., 2011).

Linear Forward Selection based feature selection:

Linear Forward Selection (LFS) algorithm is a bottom-up search procedure. It starts from an empty feature set S and gradually adds features selected by some evaluation function. The mail objective is to minimize the mean square error (MSE). The feature to be included in the feature set is selected among the remaining available features of the feature set in each iteration which have not been added to the feature set. So, the new extended features set should produce a minimum classification error compared to the addition of any other feature. LFS is widely used for its speed and simplicity. Many variants and applications have been proposed based on the LFS algorithm (A. Marcano-Cedeño *et al*., 2010).

RESULTS AND DISCUSSION

The raw dataset was subjected to the process of feature selection using a hybrid approach that combines Correlation Based Feature Selection and Linear Forward Selection. The number of genes selected using this approach is tabulated as under in Table 2

Table 2: Number of Genes in raw dataset Vs Genes selected using CFS+LFS

Dataset	Number of Genes in the raw dataset	Number of Genes selected using CFS and LFS
ALL/AML	7129	52
OVARIAN CANCER	15154	36
LUNG CANCER	12533	163

It is clearly evident from the above table that the process of feature selection has removed the redundant genes from the dataset and selected 0.7% genes in ALL/AML dataset, 0.2% genes in Ovarian Cancer dataset and 1.3% genes in Lung Cancer dataset.

The different classifiers were applied on three different datasets from Kentridge Biomedical Repository. The different cancer datasets used are ALL/AML, Ovarian Cancer and Lung Cancer. The different classifiers used are from the WEKA tool namely the Naïve Bayes, Random Forest, Bagging, Functional Tree, Lib Linear and LibSVM that were evaluated by 10-fold cross validation. The percentage of accuracy, precision, recall and f-measure of the different classifiers on the three different datasets are tabulated as under in Table3 and Table4

Table 3: Percentage of Accuracy, Precision, Recall and F-Measure of different classifiers

	NAÏVE BAYES				RANDOM FOREST				BAGGING			
	Accuracy(%)	Precision	Recall	F-Measure	Accuracy(%)	Precision	Recall	F-Measure	Accuracy(%)	Precision	Recall	F-Measure
ALL/AML	98.61	0.987	0.986	0.986	95.83	0.961	0.958	0.958	94.4	0.944	0.944	0.944
OVARIAN CANCER	100	1	1	1	98.81	0.988	0.988	0.988	98.02	0.98	0.98	0.98
LUNG CANCER	100	1	1	1	100	1	1	1	97.79	0.978	0.978	0.978

Table 4: Percentage of Accuracy, Precision, Recall and F-Measure of different classifiers

	FUNCTIONAL TREE				LIB LINEAR				LIB SVM			
	Accuracy(%)	Precision	Recall	F-Measure	Accuracy(%)	Precision	Recall	F-Measure	Accuracy(%)	Precision	Recall	F-Measure
ALL/AML	95.83	0.963	0.958	0.959	97.22	0.972	0.972	0.972	65.27	0.426	0.653	0.516
OVARIAN CANCER	100	1	1	1	100	1	1	1	100	1	1	1
LUNG CANCER	100	1	1	1	100	1	1	1	82.87	0.687	0.829	0.751

Naïve Bayes, Functional Tree and Lib Linear classifiers perform well with all the datasets. The Random Forest and Bagging Classifiers also produce comparatively high accuracy. Lib SVM does not perform well as this classifier fails for high dimensional data.

Conclusion:

In this paper, we have adopted the correlation based feature selection in combination with Linear Forward Selection to select the relevant genes and then used 6 different classifiers to evaluate the classification performance. Experimental results proved that the process of gene selection is simplified by using the adopted feature selection method. The feature selection method significantly reduces the number of genes needed for classification and has also contributed to the improvement in classifier accuracy.

REFERENCES

Ammu Prasanna Kumar and Preeja Valsala, 2013, “Feature Selection for high Dimensional DNA Microarray data using hybrid approaches”, Bioinformation., 9(16): 824-828.

Blum, A and P. Langley, 1997. “Selection of relevant features and examples in machine learning”, Journal of Artificial Intelligence, 97: 245-271.

Cheng-San Yang, Li-Yeh Chuang, Chao-Hsuan Ke, and Cheng-Hong Yang, Member, IAENG, 2008. "A Hybrid Feature Selection Method for Microarray Classification", IAENG International Journal of Computer Science.

Cosmin Lazar, Jonatan Taminau, Stijn Meganck, David Steenhoff, Alain Coletta, Colin Molter,

Virginie de Schaetzen, Robin Duque, Hugues Bersini, and Ann Nowe, 2012. “A Survey on Filter Techniques for Feature Selection in Gene Expression Microarray Analysis”, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 9(4): 1106-1119

Dash, M and H. Liu, 2000. “Feature selection for classification”, Intelligent Data Analysis., 1(1): 131-156.

Florin Leon, Catalin Lisa, Silvia Curteanu, 2010. “Prediction of the Liquid-Crystalline Property Using Different Classification Methods”, Molecular Crystals Liquid Crystals, Taylor & Francis Group, LLC, 518: 129-148

Gama, J. 2001. **“**Functional trees for classification”, Proceedings IEEE International Conference on Data Mining, ICDM pp: 147-154.

Govindarajan, M and R.M. Chandrasekaran, 2012, “Intrusion detection using neural based hybrid classification methods”, Computer Networks, 55(8): 1662-1671.

Gulshan Kumar and Krishan Kumar, 2012. “The Use of Artificial-Intelligence-Based Ensembles for Intrusion Detection: A Review”, Applied Computational Intelligence and Soft Computing, 2012: 1-20.

Guyon, I and A. Elisseeff, 2003, “An introduction to variable and feature selection,” The Journal of Machine Learning Research., 3: 1157-1182.

Harry Hochheiser and Ben Shneiderman, 2001. **“**Interactive Exploration of Time Series Data”, Lecture Notes in Computer Science, Springer Verlag, 2226: 441-446.

Jian-Hua Huang, Rui-Hua He, Lun-Zhao Yi, Hua-Lin Xie,Dong-sheng Cao,Yi-Zeng Liang, 2013. “Exploring the relationship between 5′AMP-activated protein kinase and markers related to type 2 diabetes mellitus” , Talanta, Elsevier, 110: 1-7.

Joao Gama, 2004. Functional Trees, Machine Learning, 55: 219-250.

Joaquim, F., Pinto da Costa, Hugo Alonso and Lui´s Roque, 2011. “A Weighted Principal Component Analysis and Its Application to Gene Expression Data”, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 8(1): 246-252.

Kohavi, R and G.H. John, 1997. “Wrappers for feature subset selection”, Journal of Artificial Intelligence, 97(1-2): 273-324.

Kumaravel, A and A. Aarthi, 2013. “Malware Classification based on Clustering and classification”, International Journal of Advanced Research in Computer Science and Software Engineering, 3(5): 121-123.

Li, M., Fu and Eun Seog Youn, 2003. “Improving Reliability of Gene Selection From Microarray Functional Genomics Data”, IEEE Transactions on Information Technology in Biomedicine, 7(3): 191-196.

Li-Yeh Chuang,Cheng-HueiYang,Kuo-ChuanWu,Cheng-HongYang, 2011. "A hybrid feature selection method for DNA microarray data",Computers in Biology and Medicine, Elsevier, pp: 228-237.

Marcano-Cedeño, A., J. Quintanilla-Domínguez, M.G. Cortina-Januchs and D. Andina, 2010. “Feature Selection Using Sequential Forward Selection and classification applying Artificial Metaplasticity Neural Network”, IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society, IEEE, pp: 2845- 2850.

Mark, A. Hall, 1999. “Correlation-based Feature Selection for Machine Learning”, PhD Thesis, University of Waikato

Pradipta Maji, 2012. “Mutual Information-Based Supervised Attribute Clustering for Microarray Sample Classification”, IEEE Transactions on Knowledge and Data Engineering, 24: 1.

Provan, G and M. Singh, 1995,“Learning bayesian networks using feature selection”, Fifth International Workshop on Artificial Intelligence and Statistics, pp: 450-456.

Qinbao Song, Jingjie Ni, and Guangtao Wang, 2013. “A Fast Clustering-Based Feature Subset Selection Algorithm for High-Dimensional Data”, IEEE Transactions on Knowledge and Data Engineering, 25: 1.

Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh.,Xiang-Rui Wang, Chih-Jen Lin, 2008, “LIBLINEAR: A Library for Large Linear Classification”, Journal of Machine Learning Research, 9: 1871-1874.

Shadab Adam Pattekari and Asma Parveen, 2012. “Prediction system for heart disease using Naïve Bayes”, International Journal of Advanced Computer and Mathematical Sciences, 3(3): 290-294

Sina Tabakhi,Parham Moradi, Fardin Akhla ghian, 2014. "An unsupervised feature selection algorithm based on ant colony optimization", Engineering Applications of Artificial Intelligence, Elsevier, pp: 112-123.

Van ’t Veer. L.J., H. Dai, M.J. Van de Vijver, Y.D. He, A.A. Hart, M. Mao, H.L. Peterse, K. Van der Kooy, M.J. Marton, A.T. Witteveen, G.J. Schreiber, R.M. Kerkhoven, C. Roberts, P.S. Linsley, R. Bernards and S.H. Friend, 2002. “Gene expression profiling predicts clinical outcome of breast cancer”, Nature Journal, Vol. 406

Van de Vijver, M.J., Y.D. He, L.J. Van’t Veer, H. Dai, A.A. Hart, D.W. Voskuil, G.J. Schreiber, J.L. Peterse, C. Roberts, M.J. Marton, M. Parrish, D. Atsma, A. Witteveen, A. Glas, L. Delahaye, T. Van der Velde, H. Bartelink, S. Rodenhuis, E.T. Rutgers, S.H. Friend and R. Bernards, 2002, “A gene-expression signature as a predictor of survival in breast cancer”, The New England Journal of Medicine., pp: 347.

Yukyee Leung and Yeungsam Hung, 2010. "A Multiple-Filter-Multiple-Wrapper Approach to Gene Selection and Microarray Data Classification", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 71: 108-117.