Journal of educational and social research
Volume VIII Number III September 2022

An ensemble feature selection method for gene expression classification

P. Nancy Vincentina Mary
Assistant Professor, PG Department of Computer Applications, Fatima College, Madurai, Tamil Nadu&Research Scholar, Department of Computer and Information Science, Faculty of Science, Annamalai University, Chidambaram,Tamil Nadu,India.

Dr. R. Nagarajan
Assistant Professor, Department of Computer and Information Science, Faculty of Science, Annamalai University, Chidambaram,Tamil Nadu,India.

Abstract

DNA microarray technology has modernized its approach to biological research, allowing scientists to simultaneously measure the expression levels of thousands of genes in a single experiment. Gene expression profiles that represent the state of cells at the molecular level have great potential as medical diagnostic tools. However, compared to the number of genes involved, the available training datasets generally have a relatively small sample size for classification. Limitations on these training data pose challenges for certain classification methods. Feature selection techniques can be used to extract marker genes that effectively affect classification accuracy by eliminating unwanted, noisy and redundant genes. This paper describes the ensemble feature selection techniques used in cancer classification of microarray databases and the key role of SVMs in cancer classification.

Key words: 1. Microarray, 2. Feature selection, 3. Cancer classification, 4. Gene selection, 5. Filter and Wrapper Feature Selection.

1. Introduction

DNA microarrays are an excellent high-throughput technology that can simultaneously monitor the expression levels of thousands of genes. Analysis of gene expression data is one of the major topics in health informatics [1] today. For example, classification of DNA microarray data enables the discovery of hidden patterns in expression profiles, opening up the possibility of accurate cancer classification.

The biggest challenge in classifying gene expression data is the issue of curse of dimensionality. There are many genes (features) compared to a small sample size [2, 3]. To overcome this, feature selection is used to identify the genes that are differentially expressed and eliminate irrelevant genes. Gene selection is an important task in improving the accuracy and speed of the classification system [4]. Feature selection generally falls into three categories: filters, wrappers, and embedded methods. They are categorized based on how feature selection techniques are combined with the construction of classification models.

Filter Method: Filter methods are usually used as a preprocessing step. Feature selection is independent of machine learning algorithms. Instead, features are selected based on their scores on various statistical tests of correlation with the outcome variable. Filter Methods

mainly act as rankers, ordering the features from best to worst. The ranking of features depends on the essential properties of the data, for example, variance, consistency, distance, information, correlation, etc. Figure 1 shows the flow of Filter Method.

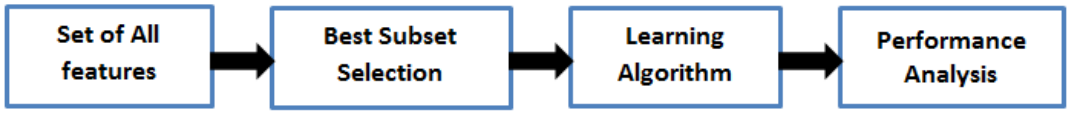

Figure 1: Filter Method

Wrapper Model: In wrapper methods, subsets of features are used to train a model. Based on the inferences from the trained model, features are added or removed from the subset. In other words, Wrapping methods compute a model with a specific subset of features and evaluate the importance of each feature. It then iterates and tries different subsets of the feature until it reaches the optimal subset. Figure 2 shows the flow of Wrapper Method.

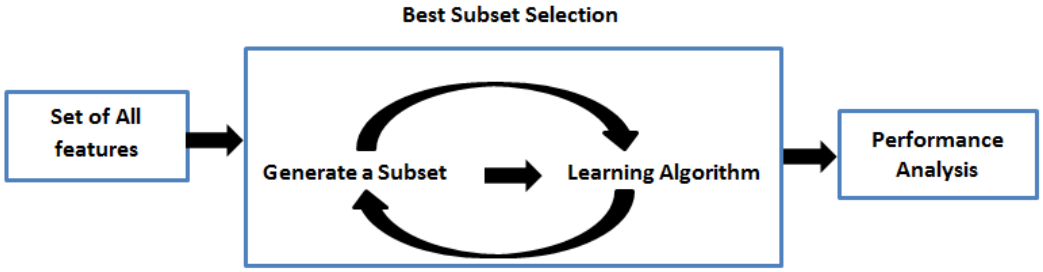

Figure 2: Wrapper Method

Embedded Method: Embedded methods bridge the gap between filters and wrappers. To begin with, this method fuse measurable and statistical criteria like a filter to choose some features, and then using a machine learning algorithm, this method pick the subset with the best classification performance. Figure 3 shows the flow of Embedded Method.

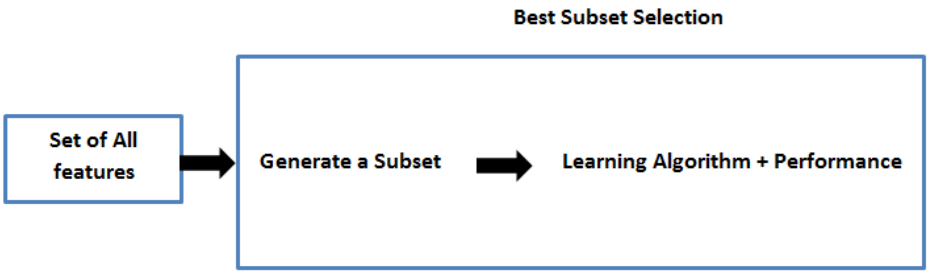

Figure 3: Embedded Method

An extensive amount of literature has been published on gene selection methods for building effective classification models. In this paper, filter and wrapper-based feature selection techniques are used to classify cancers, and also used SVM Classifier to classify cancers.

2. DNA Microarray

Microarrays provide an efficient way to collect data that can be used to determine the expression patterns of thousands of genes. The mRNA expression pattern of tissues with different normal and pathological conditions may provide information about which genes and environmental conditions can lead to disease. The experimental steps for a typical microarray began with the extraction of mRNA from a tissue sample or probe. The mRNA is then labeled with fluorescent nucleotides, eventually producing fluorescent (usually red) cDNA. The sample is later incubated with a similarly processed cDNA reference (usually green). The labeled probe and reference are then mixed and applied to the surface of the DNA microarray to attach the fluorescent sequence of the probe reference mix to the cDNA attached to the glass slide. Attracting the labeled cDNA to the probe and referencing a particular point on the microarray depends on the extent to which the sequence in the mixture (probe-reference) complements the DNA immobilized on the slide. The perfect complementarity in which the nucleotide sequence on the strand of cDNA exactly matches the sequence of DNA added to the slide is called hybridization. Hybridization is an important element of microarray technology. The assembled microarray is then laser excited and the fluorescence obtained at each point on the microarray is measured. If neither the probe nor the reference sample hybridizes to the gene on the slide, the spot will appear black. However, if hybridization occurs primarily in the probe, the spot will be red (Cy5). Conversely, when hybridization occurs primarily between the reference and the DNA immobilized on the slide, the spot fluoresces green (Cy3). If the cDNAs of the probe and reference samples hybridize equally at a particular spot, the spot may turn incandescent yellow. This indicates that certain spots share the same number of complementary nucleotides. Image processing software is used to digitize the red-to-green fluorescence and output ratio values that indicate gene expression. Figure 4 shows the process of a microarray experiment.

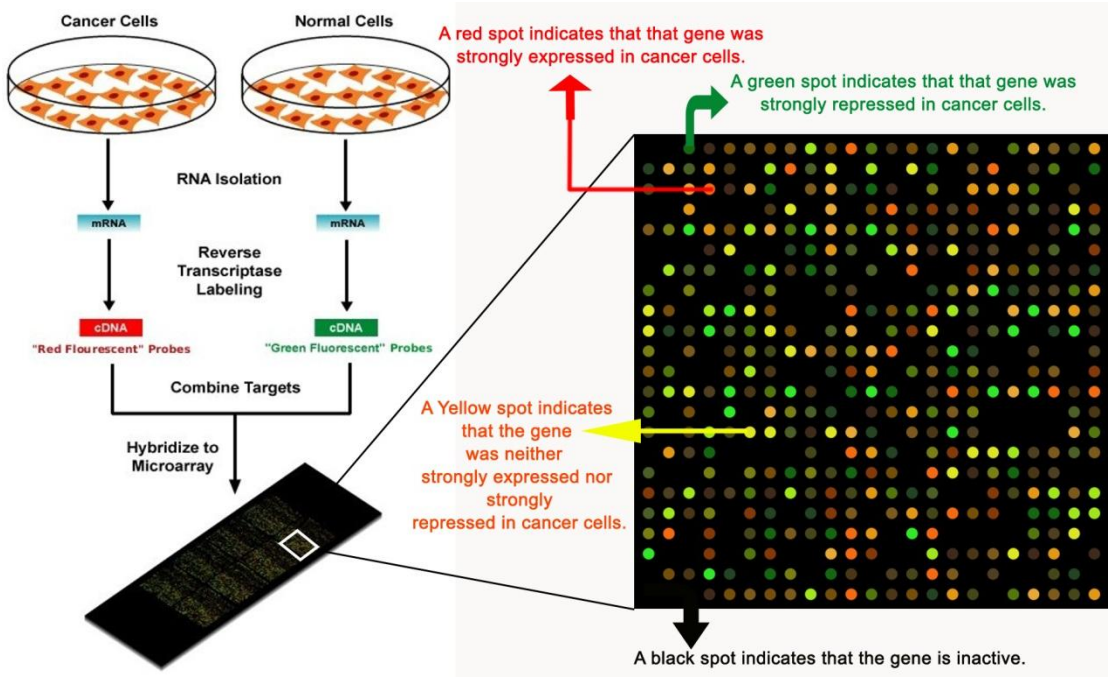

Figure 4: Microarray Experiment

Image processing software is used to digitize the red-to-green fluorescence and output ratio values that indicate gene expression. Finally, the data from all the samples are entered into

the tables that form the gene expression matrix G. The row of G corresponds to an individual gene and the column corresponds to an individual sample.

Finally, the gene expression data set can be noted by the following matrix M{w_{ij} |1≤ i≤ n, 1≤ j ≤ m}, where the rows (G { g_1 ,…,g_n }) from the expression patterns of genes, the columns (S { s_1 ,…,s_m}) from the expression profiles of samples, and w_{ij} is the measured expression level of gene i in sample j . Thus, M is defined as:

$$M = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nm} \end{bmatrix} \leftarrow g_i, i = 1, \cdots, n$$

$$\uparrow$$

$$s_j, j = 1, \cdots, m.$$

The high throughput of microarray data poses new challenges in data analysis. Although the type of analysis depends on the research questions posed, typical steps in the analysis of microarray data are as follows: i) Preprocessing and normalization, ii) detection of genes withSignificant fold changes, iii) Expression profile classification and clustering.

3. Challenges in microarray data analysis

Many microarray challenges need to be addressed before gaining new insights into gene expression. Some of the issues are:

Bias and confounding issues: Occurs during the research and design phase of microarrays and can lead to false conclusions. Technical factors such as physical differences, the number of reagents used, and differences in technician skill levels can lead to bias. Confounding, on the other hand, occurs when other factors distort the true relationship between the study variables of interest.

Cross-platform comparisons: Cross-platform comparisons of gene expression studies are difficult to perform when microarrays are constructed using different standards. Therefore, the result cannot be reproduced. To solve this problem, minimal information (MIAME) [5] on microarray experiments has been developed to improve the reproducibility, sensitivity, and robustness of gene expression analysis.

Microarray data is high dimensional data characterized by thousands of genes in few sample sizes, which cause significant problems such as irrelevant and noise genes, complexity in constructing classifiers, and missing multiple gene expression values due to inappropriate scanning. Also, moststudiesusingmicroarraydatasufferfromdataoverfittingthatrequiresadditionalvalidation.

Mislabeled dataor questioned tissue can also create other types of disadvantages for professionals, reduce the accuracy of experimental results, and lead to inaccurate conclusions about gene expression patterns.

Biological relevance results are another important criterion to consider when analyzing microarray data, rather than focusing solely on the accuracy of cancer classification. While it is undoubtedly important to obtain very accurate classification results from microarray data analysis, it is also important to disclose biological information during the cancer classification process. For example, identifying genes that are under expressed or overexpressed in cancer cells may help subject experts design and plan better treatments for cancer patients. Therefore, most subject experts are interested in classifiers that not only achieve high classification accuracy, but also reveal important biological information.

4. Feature selection techniques in micro array data analysis

DNA microarray technology is used to measure changes in gene expression levels. The expression of this genetic information takes place in her two stages, the transcription stage and the translation stage. In transcription, the DNA molecule is rewritten into mRNA, whereas in the translation stage, mRNA is translated into the amino acid sequence of the corresponding protein. DNA microarray analysis provides simultaneous access to thousands of genes by recording their expression levels simultaneously. Altered gene expression has been shown to be associated with different types of cancer. Cancer classification using gene expression data is an important task due to the nature of gene expression data. Expression data have a very high dimensionality, typically on the order of thousands to tens of thousands of genes. The situation becomes more complicated with sample size numbers typically less than 100. High feature dimensionality and small population size usually lead to overfitting of the classifier. The term Curse of Dimension was coined to refer to this situation. Computational cost also poses an important limitation. Another important issue is that not all genes are associated with cancer, making it difficult to extract biologically relevant genes.

Dimensionality reduction techniques in gene selection can be classified into two categories. They are transformation or selection-based reduction. The main difference in classification is whether the dimensionality reduction technique transforms or preserves the semantics of the records in the reduction process. Transformation-based reduction, such as principal component analysis (PCA), transforms the original features of the data set, typically reducing the number of uncorrelated features called principal components. In contrast, selection reduction techniques attempt to determine a minimal subset of features from the problem domain while preserving the importance of the original feature set. Therefore, selection-based reduction techniques are largely preferred in many bioinformatics applications, especially microarray data analysis, as they offer the advantage of subject matter expert interpretability. Feature selection is the process of systematically reducing the dimensionality of a data set to the optimal subset of attributes for classification purposes. The problem of feature selection is therefore an important problem in cancer classification. The feature selection process has been shown to improve the predictive power of classifiers in many applications.

The objectives of feature selection techniques are many, the first one is to avoid over fitting and improve model performance, and for example selecting highly informative genes could enhance the accuracy of classification model. The second one is to give faster and more cost-effective models. The third one is to gain a deeper insight into the underlying processes that generated the data. Even though, feature selection techniques have many benefits, it also introduces extra complexity level which requires thoughtful experiment design to address the challenging tasks, yet provide fruitful results. Feature selection techniques can be organized into three categories, such as filter method, wrapper method and embedded method based on how they combine the feature selection search with the construction of the classification model.

The filtering method ranks each feature by a unique metric, and only the highest-rated features are used while the remaining low-ranked features are discarded. This method also relies on the general characteristics of the training data to select certain features without involving any learning algorithm. Therefore, the results of the filter model will not affect any of the classification algorithms. Furthermore, the filtering methods also provide a very easy computation and can be simply adapted to large-scale microarray datasets because they have short runtimes. Univariate filtering methods such as Bayesian networks [6], information gain (IG) and signal-to-ratio (SNR) [7] and Euclidean distance [8] have been widely used in microarray data to identify informative genes. Information Gain has been reported to be the superior gene selection technique, but different types of univariate techniques seem to make sense when trained on different data sets. On the other hand, the Bayesian network seems to be the ideal

platform for integrating heterogeneous information sources. Besides applying parametric techniques to identify informative genes from microarray data [9], non-parametric techniques such as number of misclassification thresholds or TNoM scores were applied. Essentially, this technique splits the information gene by assigning it a threshold value. However, it is difficult to determine the most appropriate threshold. Other non-parametric techniques such as Pearson's correlation coefficient [10] and Significant Analysis of Microarray (SAM) [11] have been reported as the main feature selection techniques. Univariate filtering method has been widely used in microarray data analysis. This trend can be explained by a number of reasons, for example the results provided by univariate gene rankings are intuitive and easy to understand. Simplified versions of these results can meet the goals and expectations of biological and molecular experts who require confirmation of results using laboratory techniques. Furthermore, filtering methods also provide less computational time to produce results, which is an addition favored by experts in the field. However, gene ranking based on univariate method has some limitations. The main problem is that selected genes are more likely to be redundant. This means that highly ranked genes can carry similar discriminatory information for a given class. Removing a high-ranking gene may not reduce classification accuracy. Since univariate filtering methods do not count relationships between genes, an optimal gene selection method called Markov Blanket Filtering [12] has been developed, which can remove redundant genes for elimination. Based on this approach, the Redundancy Based Filter (RBF) [13] method has been proposed to solve the redundancy problems and the results are quite promising.

While the filter techniques handle the identification of genes independently, a wrapper method [14] on the other hand, inserts a gene selection method within a classification algorithm. In the wrapper methods a search is conducted in the space of genes, assessing the goodness of each found gene subset by the approximation of the accuracy ratio of the specific classifier to be used, after that the classifier is trained only with the found genes. However, the wrapper approach is claimed to produce better prediction accuracy estimates than the filter approach. Its computational complexity must be taken into account. Wrapper methods can be divided into different groups, deterministic and randomized search algorithms. A genetic algorithm (GA) is a randomized search algorithm and optimization that mimics evolution and natural genetics. It has been used for binary and multiclass cancer discrimination [15]. A common drawback of wrapper methods like GA is that they have a higher risk of overfitting than filtering techniques and are very computationally expensive. In contrast, wrapper methods involve interactions between gene selection and classification models, making them unique compared to filtering techniques.

A third class of feature selection approaches is the embedded method. The difference between the embedded method and other feature selection methods is that the search mechanism is built into the classifier model. Therefore, embedded methods are the same as wrapper methods and are specific to a particular learning algorithm. Embedded methods have the advantage of being much less computationally intensive than wrapper methods while involving interaction with the classification model. The Support Vector Machine (SVM) method of Recursive Feature Elimination (RFE) [16] was used for gene selection.

5. Ensemble filter and wrapper feature selection method

This paper combines the filter and wrapper model to select feature genes in microarrays. The methods are ReliefF filter and Wrapper. These methods are applied using Weka machine learning tool. Combining filter and wrapper is based on ensemble model. First ReliefF Filter method is applied to the dataset, after that Wrapper method is applied to the output of the Filter Method. By using ensemble model the classification accuracy can be improved.

In the first stage, the ReliefF is employed on the dataset to acquire a candidate gene set. The genes are evaluated and organized according to the ReliefF measure. Then, the ReliefF examines and chooses the top P genes, which will become the candidate gene set. This procedure filters the insignificant genes and minimizes the computational complexity for the next stage. In

the second stage, WrapperSubsetEval is used for feature selection using greedy stepwise search in this work. The perceptive behind this method is that it can calculate the importance of each feature related to the class. It is a well-known form of wrapper feature selections search method which involves adding or removing the features based on their discriminative powers.

Weka is used to determine the information value of each feature and sort the features in accordance with their information gain value. Higher values indicate higher discrimination of this feature from other categories, and mean that this feature can be used to calculate classification results effectively. After calculating the information gain values of all features, threshold is introduced. Since after calculation most information gain values were zero, not many features have an influence on the categorization of a data set. The threshold in this study is 0 for most data sets. If the information gain values of the features are higher than the threshold, feature is selected, if not, the feature is not selected. Figure 5 describes Ensemble filter and wrapper model feature selection method.

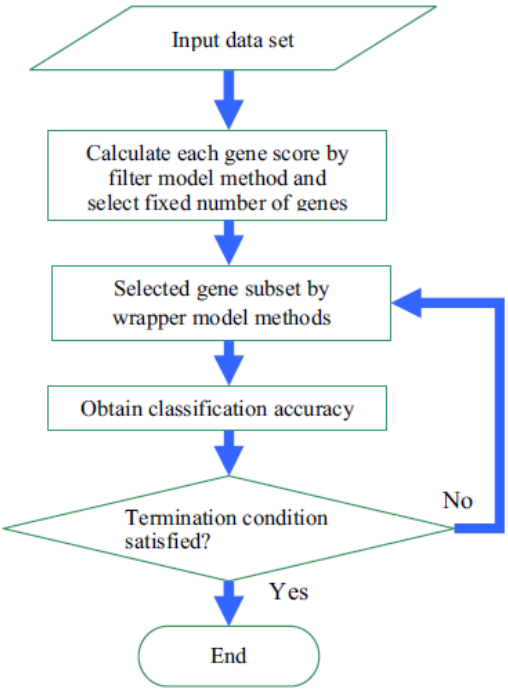

Figure 5: Ensemble filter and wrapper model feature selection method

During the features selection process, the variables that still have a non-zero coefficient after the shrinking process are selected to be part of the model, with the goal of minimizing the prediction error. Prediction analysis is carried out on all the datasets using classifiers which are commonly used in machine learning predictive analysis. In this paper Support Vector Machine is used. This classifier is evaluated based on the accuracy, Sensitivity and Specificity both before and after feature selection was conducted on the datasets.

6. Support Vector Machine for classification (SVM)

Gene expression Microarrays are very beneficial for clinical decision support system in the form of prediction and diagnosis of clinical results of cancer and other complex diseases. To increase the benefits of this technology, researchers are continuously working to develop and apply the most accurate decision support algorithms for patient gene expression profiling.Prior research suggests that among well-established and popular techniques for classification of microarray gene expression data, support vector machine (SVMs) achieve the best classification

performance, significantly out performing back propagation neural networks, weighted voting methods, probabilistic neural networks,K nearest neighbors and decision trees.The reasons for this are

1. SVMs have demonstrated the ability to not only correctly separate entities into appropriate classes, but also to identify instances whose established classification is not supports by the data.
2. SVM has many attractive mathematical features for gene expression analysis. They arethe ability to handle large feature spaces, the ability to identify outliers,their flexibility in choosing a similarity function and sparseness of solution when dealing with large data sets.

Supervised classification, also called prediction or discrimination, involves developing algorithms for previously defined categories. Algorithms are usually developed on a training dataset and then tested on an independent test dataset to estimate the accuracy of algorithms. Support vector machine (SVM) is a supervised Machine learning algorithm used for classification, regression and outlier analysis. The simplest type of support vector machines is linear classification which attempts to draw a straight line that separates data by two dimensions. Many linear classifiers (also called hyperplane) are able to separate the data. However, only one achieves maximum separation. In 1963, Vapnik proposed a linear classifier as an original optimal hyper plane algorithm. Replacing the dot product by a nonlinear kernel function allows the algorithm to fit the maximum-margin hyper plane in the transformed feature space. The SVM finds linearly separable hyperplane with maximal margin in this high-dimensional space, called the kernel function.

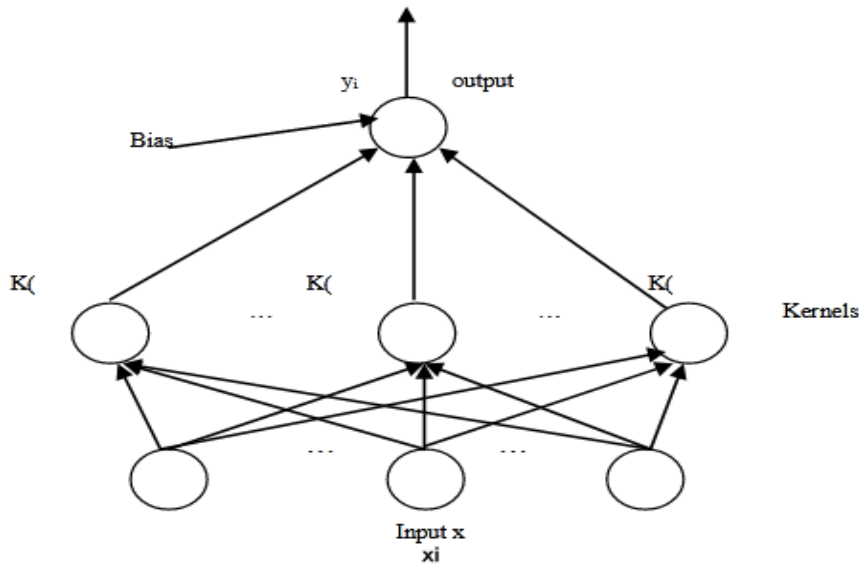

Figure 6: Structure of an SVM

In Figure 6, the notation xi denotes the i^{th} vector in a dataset {(xi, yi)}, i = 1 to n where yi is the label associated with xi. Object xi is also called pattern, input, and example.The K from i = 1 to n represents the kernel functions of real-valued data. Given a training set of instance-label pairs and Y the support vector machine (SVM) need the solution of the subsequent optimization problem:

Here, the training vectors are mapped into a higher dimensional space by the function. The simplest form of a prediction problem is binary classification: trying to differentiate between objects that belong to one of two groups, positive (+1) or negative (-1). SVMs use two key models to solve this problem: kernel functions and large-margin separation.

SVM first maps the input to a high-dimensional feature space and finds a separating hyperplane in that space that maximizes the margin between the two classes.Maximizing the margin is a quadratic programming (QP) problem, which can be solved from the dual problem by introducing Lagrangian multipliers. Without any information of the mapping, the SVM finds the optimum hyperplane by using the dot product functions in feature space, and that are called kernels. An optimal hyperplane solution can be described as a combination of several input points called support vectors.[17]. SVM fit in to the general category of kernel methods. The kernel method is a data-dependent algorithm with dot-products.In this case the dot product can be replaced by a kernel function that computes the dot product in a high-dimensional feature space.

The four basic kernels are Linear, Quadratic, Polynomial and Radial Basis Function (RBF). The kernel function defines the feature space in which the training set examples are classified, so choosing a good kernel function is important.SVM classifier is used in this work, because of its high accuracy, capability to deal with high-dimensional data such as gene expression, and flexibility in exhibitingvarious sources of data [18].

7. Results and discussion

The dataset of Breast cancer and ovarian cancer are originally presented by Zexuan Zhu, Y. S. Ong and M. Dash [19]. Table 1, describes the dataset characteristics. Breast cancer dataset have two class labels as relapse and non- relapse, Ovarian cancer dataset have two class labels as Cancer and normal.

Table – 1
Dataset characteristics

Dataset	Gene	Instances	Class
Breast Cancer	24481	97	2
Ovarian Cancer	15154	253	2

Ensemble filter and wrapper feature selection method's performance on the classification of breast cancer and ovarian cancer datasets are tested. The accuracy of the classifier with regard to the original high and low dimensional datasets is evaluated.

First without feature selection, classification is done on high dimensionality data. Bayes Net, Naïve Bayes, k-NN and Support Vector Machine classifiers are used. A 10-fold cross validation is applied with each classifier for the training and testing. The result in Table 2 illustrates, Bayes Net, Naïve Bayes, k-NN and Support Vector Machine's classification accuracy.

Table – 2
Accuracy of Classifiers with Original Dataset

Dataset	Classifier			
	Bayes Net	Naïve Bayes	k-NN	SVM
Breast Cancer	86.71	83.87	77.58	**88.71**
Ovarian Cancer	82.35	62.75	85.29	**91.18**

Next, Ensemble filter and wrapper feature selection is applied to the dataset, and then classifiers are used to classify the dataset. The accuracy of the classifiers improved after applying the feature selection method. Table 3 illustrates, Bayes Net, Naïve Bayes, k-NN and Support Vector Machine's classification accuracy after applying feature selection method.

Table – 3
Accuracy of Classifiers after applying ensemble feature selection method

Dataset	Classifier			
	Bayes Net	**Naïve Bayes**	**k-NN**	**SVM**
Breast Cancer	87.12	84.01	79.34	**90.65**
Ovarian Cancer	84.78	65.79	86.34	**93.21**

Figure 7 Shows that, SVM classifier is more accurate than Bayes Net, Naïve Bayes and k-NN classifiers in the breast cancer dataset.

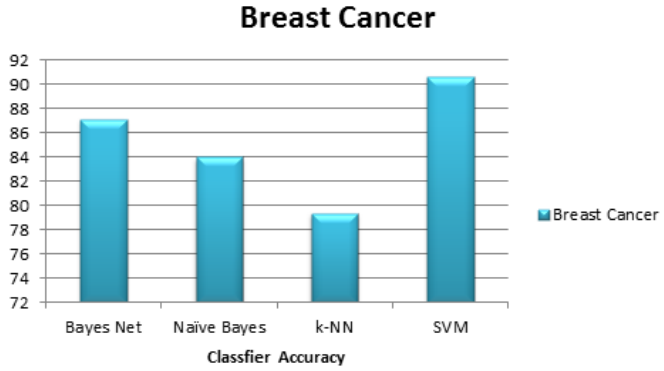

Figure 7: SVM Classifier accuracy in Breast cancer dataset

Figure 8 Shows that, SVM classifier is more accurate than Bayes Net, Naïve Bayes and k-NN classifiers in the ovarian cancer dataset.

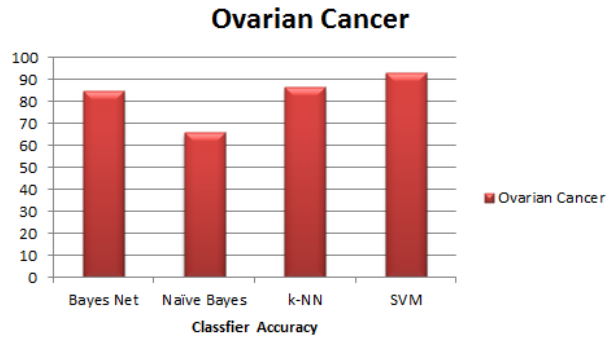

Figure 8: SVM Classifier accuracy in Ovarian cancer dataset

The experimental results show that the accuracy of microarray data classification which has feature selection is better than without feature selection.

9. Conclusion and future directions

In this paper, ensemble filter and wrapper model based feature selection for microarray classification is used. Then, Bayes Net, Naïve Bayes, k-NN and Support Vector Machine are used to evaluate the classification performance. Experimental results showed that the proposed method simplified gene selection and the total number of parameters needed effectively, thereby obtaining a higher classification accuracy compared to other feature selection methods. The classification accuracy obtained by the proposed method was higher than other methods. In the future, other available or modified filter and wrapper based feature selection can be integrated and tested with other available or modified supervised classifiers.

References

1. Ng, R. T., & Pei, J. (2007). Introduction to the special issue on data mining for health informatics. ACM SIGKDD Explorations Newsletter, 9(1), 1-2.
2. Piatetsky-Shapiro, G., & Tamayo, P. (2003). Microarray data mining: facing the challenges. ACM SIGKDD Explorations Newsletter, 5(2), 1-5.
3. Rocha, M., Mendes, R., Maia, P., Glez-Peña, D., &Fdez-Riverola, F. (2007, July). A platform for the selection of genes in DNA microarray data using evolutionary algorithms. In Proceedings of the 9th annual conference on Genetic and evolutionary computation (pp. 415-423).
4. Shang, C., &Shen, Q. (2005). Aiding classification of gene expression data with feature selection: a comparative study. International Journal of Computational Intelligence Research, 1(1), 68-76.
5. Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., ...&Vingron, M. (2001). Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. Nature genetics, 29(4), 365-371.
6. Giallourakis, C., Henson, C., Reich, M., Xie, X., &Mootha, V. K. (2005). Disease gene discovery through integrative genomics. Annual review of genomics and human genetics, 6(1), 381.
7. Wang, Z. (2005). Neuro-fuzzy modeling for microarray cancer gene expression data. First year transfer report, University of Oxford, 24.
8. Cho, S. B., & Won, H. H. (2003, January). Machine learning in DNA microarray analysis for cancer classification. In Proceedings of the First Asia-Pacific Bioinformatics Conference on Bioinformatics 2003-Volume 19 (pp. 189-198).
9. Ben-Dor, A., Bruhn, L., Friedman, N., Nachman, I., Schummer, M., &Yakhini, Z. (2000, April). Tissue classification with gene expression profiles. In Proceedings of the fourth annual international conference on Computational molecular biology (pp. 54-64).
10. Hu, H., Li, J., Wang, H., &Daggard, G. (2006, October). Combined gene selection methods for microarray data analysis. In International conference on knowledge-based and intelligent information and engineering systems (pp. 976-983). Springer, Berlin, Heidelberg.
11. Fung, B. Y., & Ng, V. T. (2003). Classification of heterogeneous gene expression data. ACM SIGKDD Explorations Newsletter, 5(2), 69-78.
12. Koller, D., &Sahami, M. (1996). Toward optimal feature selection. Stanford InfoLab.
13. Yu, L., & Liu, H. (2004, August). Redundancy based feature selection for microarray data. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 737-742).
14. Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. Artificial intelligence, 97(1-2), 273-324.
15. Liu, J. J., Cutler, G., Li, W., Pan, Z., Peng, S., Hoey, T., ...& Ling, X. B. (2005). Multiclass cancer classification and biomarker discovery using GA-based algorithms. Bioinformatics, 21(11), 2691-2697.
16. Bhola, A., &Tiwari, A. K. (2015). Machine learning based approaches for cancer classification using gene expression data. Machine Learning and Applications: An International Journal (MLAIJ), 2(3/4).
17. Lin, C. F., & Wang, S. D. (2002). Fuzzy support vector machines. IEEE transactions on neural networks, 13(2), 464-471.
18. Schölkopf, B., Tsuda, K., &Vert, J. P. (2004). Kernel methods in computational biology. MIT press.
19. Zhu, Z., Ong, Y. S., & Dash, M. (2007). Markov blanket-embedded genetic algorithm for gene selection. Pattern Recognition, 40(11), 3236-3248.