Hybrid Neuro Fuzzy Runge-Kutta (HNFRK) classifier based Gene Selection for Cancer Classification on Microarray Gene Expression Data

K.Sathishkumar[1], Dr.V.Thiagarasu[2], Dr.E.Balamurugan[3], Sujith Jayaprakash[4]
BlueCrest College, Ghana, West Africa[1, 3, 4]
Gobi Arts & Science College (Autonomous), Tamilnadu, India[2]

Abstract: This paper proposes a gene selection framework, based on hybrid neuro-fuzzy approach for cancer classification. Gene selection is often used in preprocessing microarray data for efficient classification of diseases. Gene selection approaches mostly uses filter techniques as they provide efficient and effective results. But, it is a fact that certain genes discarded by several existing techniques are helpful for classification at certain conditions and cannot be removed blindly. In this paper, a Structured Sparse Principal Component Analysis (SSPCA) approach is proposed to select the gene for classification. The SSPCA is applied to three microarray datasets to select feature genes with Hybrid Neuro Fuzzy Runge-Kutta (HNFRK) classifier to evaluate its performance. This method is compared with several previous gene selection results to show that this SSPCA gene selection algorithm has good classification accuracy and model stability. HNFRK gives the best results for original data of all the datasets and the predictions for noisy data are adequate in comparison with three others classifiers. HNFRK is best for less number genes, clearly when compared to conventional ANFIS.

Keywords: HNFRK; Gene Selection; Microarray Gene Expression Data Analysis; Cancer Classification

1 Introduction

Microarray (also known as biochip) technology can provide many large datasets of biomarkers for biologists. There are many possible machine learning methods for microarrays data analysis. Identification of functions of cells is an important challenge in cancer diagnosis. There are different types of microarrays for biological studies. Some of microarrays are made of tumors and normal tissues. Investigation of patterns of gene activities for mentioned microarrays can lead biologists to find better solutions for prognosis cancers. Early predictions of cancer increase the chance of disease improvement and decrease cost of medical care for patients. Besides, finding related biomarkers helps biologists to predict cancer disease effectively. So we can define a cancer diagnosis framework as a classification problem. In classification problems, there is a pair of a vector of features and a label for every record in datasets. After a training phase with train data, the classifiers such as Support Vector Machine(SVM), K-Nearest Neighbor(KNN), Classification And Regression Trees

(CART), would map some of input features to one of the existing labels. Microarrays usually have many biomarkers as same as features in datasets. Every biomarker shows an expression level of a gene. Microarrays have huge redundancy and high dimensionality, hence feature selection or in other words, gene selection is a main phase of microarray samples classification for cancer prognosis. In general, feature selection methods can be divided into two categories. In the first category based on filtering[1], feature selection and classification method are proceeded distinctively. A single or multiple selection criteria must satisfy to identify a final subset of features. Because of independence between two phases, filter model is fast but it needs to select exacting methods and limitations for achieving high precision. In the second category is called wrapped model[2], feature selection process must be embedded for each classifier and precision is achieved overall. This approach needs more computation than the filtering method but it is possible to achieve better precision because of optimized feature selection for particular classifier. A classifier should be run with different groups of candidate features and after evaluating the result, those which are more efficient would be selected. There are many published papers about the gene selection in cancer diagnosis which using different approaches and various classifiers, so good results were obtained from mentioned classifiers under different gene selection methods and a large number of genes. However, this issue was less considered by researchers in related works that microarrays due to their essential features have noisy nature [3]. Because of different environmental conditions and tools in various biological laboratories, varying degrees of noise influences microarrays contained data generally. Final data are extracted from the preprocessed image of scanned spots. There are very bottlenecks in preprocessing such as gridding [4], noise reduction, image enhancement [5], robust spots finding in distortional images of microarray [6] and so on. Nevertheless researchers often deal with microarrays as certain data. A noisy microarray is shown in *Figure 1:*

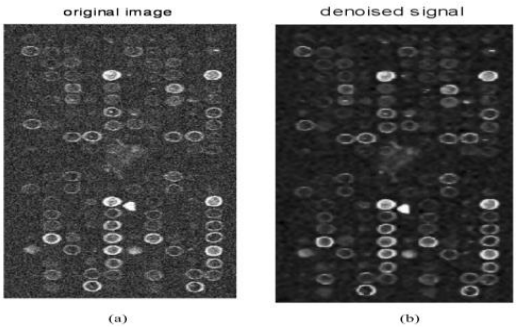

Figure 1:Original microarray image (a) Denoised microarray image(b)[10].

Due to immense growth and development of microarray technologies, it is possible to quantify the expression levels of millions of genes in a single experiment [2]. It ultimately constructs microarray data. A comparison analysis between the gene expression levels of abnormal and normal tissues can also be done. This comparison is very essential and useful to choose genes that might expect the clinical behavior of cancers. Therefore, there is a requirement to choose informative genes that contribute to an abnormal state. But, the gene selection process is a challenging task mainly due

to the features of microarray data such as higher-dimensional data, irrelevant genes, and noisy data.

In order to overcome the above drawbacks, a gene selection technique is generally used to choose a subset of genes that enhances the ability of the classifier to classify samples more accurately. Novel and advanced gene selection techniques can offer a more compact gene subset without much degradation in classification accuracy. Advanced techniques can also minimize the dimensionality of data, and eliminate irrelevant and noisy genes. Moreover, a smaller number of chosen genes can be more efficiently used for diagnostic purposes in clinical settings [7].

A number of dimensionality reduction techniques have been developed and used in literature to solve the microarray gene expression classification problem. Principal Component Analysis (PCA) is one of the most widely used dimensionality reduction algorithms for the efficient analysis of gene data [8]. But, one of the major limitations of PCA is that, the factors are linear combinations of all the original variables, so most factor coefficients are not zero, which means that it is difficult to understand and visualize the PCA model with all the original variables. In order to overcome the above said limitation, various proposals have been introduced in the literature. In this paper a Structured Sparse Principal Component Analysis (SSPCA) [9] approach is proposed for selecting the required genes for the classification purpose which is processed by Hybrid Neuro Fuzzy Runge-Kutta (HNFRK) classifier.

2 Related Work

In recent years, the gene expression data has been highly focused in a nonlinear subspace. In such scenarios, the linear subspace based dimensionality reduction techniques donot provide significant results. In recent years, manifold learning-based dimensionality reduction approaches such as Isomap[10], locally linear embedding(LLE)[11] and local preserving projections(LPP)[12] have become an attractive area of research. It is believed that these techniques are effluent in identifying the intrinsic geometrical structure of the nonlinear data. The manifold learning-based dimensionality reduction techniques have also been used to deal with the microarray gene expression classification problem [13], [14]. Even though manifold learning-based dimensionality reduction techniques provide significant results, they cannot be easily applied in certain applications because of their complexity and storage necessities. More recently, sparse representation obtained by solving an optimization problem has been widely used in applications such as pattern recognition, computer vision, etc [15], [16].

Sparse representation which can represent the high dimensional data in a more effective way has been productively used in pattern recognition problems. But, it doesn't take into account, the label information of data samples. In order to overcome this drawback, Chunming Xu [17] presented a novel dimensionality reduction algorithm called Discriminatively Regularized Sparse Subspace Learning (DR-SSL). DR-SSL algorithm makes use of the sparse representation to model the data and moreover, it can effectively utilize the label information to direct the dimensionality reduction process. Moreover, this algorithm can effectively deal with the out-of-sample problem. The experiments on gene-expression data sets indicate that it is an effective tool for dimensionality reduction and gene-expression data classification.

In this paper, we use sparse principal component analysis (PCA) to solve clustering and feature selection problems. Sparse PCA seeks sparse factors, or linear combinations of the data variables, explaining a maximum amount of variance in the data while having only a limited number of nonzero coefficients. PCA is often used as a simple clustering technique and sparse factors allow us here to interpret the clusters in terms of a reduced set of variables. We begin with a brief introduction and motivation on sparse PCA and detail our implementation of the algorithm in d'Aspremont et al. (2005). We finish by describing the application of sparse PCA to clustering and by a brief description of DSPCA, the numerical package used in these experiments [18].

Yanwei Huang, Liqing Zhang [19] presented a gene selection algorithm using Multiple Principal Component Analysis with Sparsity (MSPCA). The MSPCA algorithm is used to analyze normal and disease gene expression samples and to set these component loadings to zero if they are smaller than a threshold for sparse solutions. Next, genes with zero loadings across all samples (both normal and disease) are removed before extracting feature genes. Feature genes are genes that contribute differentially to variations in normal and disease samples and, thus, can be used for classification. The MSPCA is applied to three microarray datasets to choose feature genes with a linear support vector machine to evaluate its performance. This technique is compared with a number of previous gene selection results to indicate that this MSPCA gene selection approach has good classification accuracy and model stability when compared with the previous works [19].

3 Methodology

Principal component analysis (PCA) is an important tool for data investigation and unsupervised dimensionality reduction between linear combinations of the data variables. The main drawbacks of PCA are it finds a small number of significant factors, the factor by itself normally occupy all original variables. To overcome the drawbacks several alternatives to PCA are investigated by the researchers like nonnegative matrix factorization (NMF) and sparse PCA (SPCA) etc.

3.1 Structured Sparse PCA(SSPCA)

The structured sparse PCA (SSPCA) is based on a structured condition of having been made regular is investigated by the author in [20]. Whereas classical sparse priors simply covenant with cardinality, the regularization encodes higher-order information about the data. In this paper, propose an efficient and simple optimization procedure to select the genes for classification. Let a dataset containing two classes, normal and cancer classes. Though, in several applications only confine the size of the factors does not look like accurate because the measured factors are not only predictable to be sparse but also to have a convinced structure. Indeed, the status of NMF for face image analysis is obliged to the fact that the method happens to get back the sets of variables that are localized on the face and capture some parts of the face which look like instinctively consequential given to prior information. As a result, gain the

quality of the factors persuade by enforcing directly this a priori in the matrix factorization constraints. It is to be attractive to encode higher-order information about the supports that replicates the structure of the data. The author proposed a structured sparse PCA (SSPCA), which explains the variance of the data by factors that are not only sparse but also respect some a priori structural constraints reason relevant to model the data at hand. The below section describes the detailed procedures for using the SSPCA to identify feature genes that can be used for classification.

3.2 Solution for Gene Reduction

Let a matrix $X \in \mathbb{R}^{n \times p}$ of n rows related to n observations in $\mathbb{R}^p$, the dictionary learning problem is to locate a matrix $V \in \mathbb{R}^{p \times r}$, called the *dictionary*, similarly each observation is attained by a linear combination of the r columns $(V^k)k \in \{1, \ldots, r\}$ of V called the *dictionary elements*. If $U \in \mathbb{R}^{n \times r}$ is the matrix of the linear combination coefficients or *decomposition coefficients*, the matrix product UV^T is called a decomposition of X.

The Learning parallel to the dictionary V and the decomposition U leads to a matrix factorization problem. Like in [14], it is natural when learning a decomposition, to castigate few or some norms of U and V, say Ω_u and Ω_v in that order, to encode former information sparsity about the decomposition of X. This can be as follows

$$\min_{U \in \mathbb{R}^{n \times r}, V \in \mathbb{R}^{p \times r}} \frac{1}{2np} \|X - UV^T\|_F^2 + \lambda \sum_{k=1}^{r} \Omega_v(V^k) \qquad s.t. \quad \forall k, \Omega_u(U^k) \leq 1, \quad (1)$$

where the regularization parameter $\lambda \geq 0$ controls which point the dictionary is to be regularized. Let assume that both regularizations Ω_u and Ω_v are convex, problem (1) is convex based on U for V fixed. It is still not jointly convex in (U, V) it denotes the reduced data sets.

3.2 Feature Genes by SSPCA

SSPCA is processed on the normal and cancer samples genes by means of above method gains a reduced set of genes pointing out a major variations in the normal samples called Normal Genes (NG), and a reduced set of genes representing major variations in the cancer samples called Cancer Genes(CG). Feature Genes (FG) can be firmed by

$$FG = \overline{NG \cap CG}$$

In [21] present a stop rule for the smallest number of FG as the number of principal components. At this point, if q^N and q^C are distinct as the number of principal

components for the normal and cancer samples in Eq. (2), the number of FG is in the range of $|q^N - q^C| \; to \; |q^N + q^C|$.

3.3 Classification of Genes using Hybrid Neuro Fuzzy Runge-Kutta (HNFRK) classifier

ANFIS is one of hybrid neuro-fuzzy inference expert systems and it works in Takagi-Sugeno-type fuzzy inference system, which was developed by Jang [22]. ANFIS has a similar structure to a multilayer feed forward neural network but the links in an ANFIS only indicate the flow direction of signals between nodes and no weights are associated with the links.

A. Architecture of ANFIS

The ANFIS is a fuzzy Sugeno model put in the structure of adaptive systems to make easy learning and adaptation Jang (1993). Such structure makes the ANFIS modeling more efficient and less reliant on expert knowledge. To present the ANFIS structural design, two fuzzy if–then rules based on a first order Sugeno model are measured

Rule 1: If (x is A1) and (y is B1) then (f1 = p1x + q1y + r1).

Rule 2: If (x is A2) and (y is B2) then (f2 = p2x + q2y + r2),

where x and y are the inputs, Ai and Bi are the fuzzy sets, fi are the outputs within the fuzzy region precise by the fuzzy rule, pi; qi and ri are the design parameters that are determined throughout the training process. The ANFIS architecture to put into practice these two rules is shown in Fig. 5.6, in which a circle indicates a fixed node, while a square indicates an adaptive node.

In the first layer, all the nodes are adaptive nodes. The outputs of layer 1 are the fuzzy membership grade of the inputs, which are given by

$$O\_i\hat{}1=\mu\_Ai\ (x), \qquad i=1,2,$$

$$O\_i\hat{}1=\mu\_(Bi-2)\ (y), \qquad i=3,4,$$

where μ_Ai (x), μ_(Bi-2) (y)can adopt any fuzzy membership function. For example, if the bell shaped membership function is employed, μ_Ai (x) is given by

$$\mu\_Ai\ (x)=1/\ 〖1+\{((x-c\_(i\ ))/a\_i\ )\hat{}2\}\hat{}(b\_i\ )$$

where a_i,b_i and c_(i)are the parameters of the membership function, governing the bell-shaped functions, as a result.

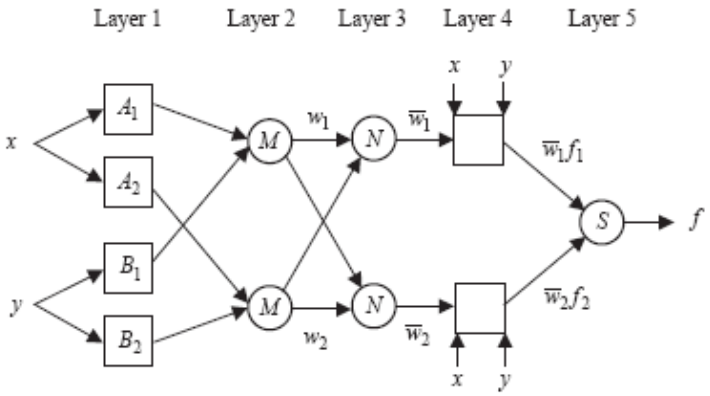

Figure 2: ANFIS architecture

In the second layer, the nodes are fixed nodes. They are labeled with M, indicating that they carry out as a simple multiplier. The outputs of this layer can be correspond to as

$$O_i^2 = w_i = \mu_{Ai}(x)\mu_{Bi}(y), \qquad i = 1,2, \quad (4)$$

which are the called as firing strengths of the rules.

In the third layer, the nodes are also fixed nodes. They are labeled with N, indicating that they play a normalization role to the firing strengths from the preceding layer. The outputs of this layer can be represented as

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \qquad i = 1,2, \qquad (5)$$

which are the so-called normalized ring strengths.

In the fourth layer, the nodes are adaptive nodes. The output of each node in this layer is simply the product of the normalized firing strength and a first-order polynomial. Thus, the outputs of this layer are given by

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i(p_i x + q_i y + r_i), \qquad i = 1,2 \qquad (6)$$

In the fifth layer, there is only one single fixed node labeled with S. This node performs the summation of all incoming signals. Hence, the overall output of the model is given by

$$O_i^5 = \sum_{i=1}^{2} \overline{w}_i f_i = \frac{(\sum_{i=1}^{2} w_i f_i)}{w_1 + w_2}, \quad i = 1,2 \qquad (7)$$

It can be experiential that there are two adaptive layers in this ANFIS structural design, that is the first layer and the fourth layer. In the first layer, there are three changeable parameters $\{a_i, b_i, c_i\}$ which are connected to the input membership functions. These parameters are the so-called premise parameters. In the fourth layer, there are also three modifiable parameters $\{p_i, q_i, r_i\}$, pertaining to the first order polynomial. These parameters are so-called consequent parameters.

B. Learning algorithm of ANFIS

The task of the learning algorithm for this architecture is to alter all the modifiable parameters, namely a_i,b_i,c_i} and 〖{p〗_i,q_i,r_i}, to make the ANFIS output match the training data. When the premise parameters a_i,b_i and c_i of the membership function are fixed, the output of the ANFIS model can be defined as

$$f = \frac{w_1}{w_1+w_2} f_1 + \frac{w_2}{w_1+w_2} f_2 \qquad (8)$$

Substituting Eq. (5) into Eq. (8) yields

$$f = \bar{w}_1 f_1 + \bar{w}_2 f_2 \qquad (9)$$

By substituting the fuzzy if–then rules into Eq. (9), it becomes

$$f = \bar{w}_1(p_1 x + q_1 y + r_1) + \bar{w}_2(p_2 x + q_2 y + r_2) \qquad (10)$$

After rearrangement, the output can be expressed as

$$f = (\bar{w}_1 x)p_1 + (\bar{w}_1 y)q_1 + (\bar{w}_1)r_1 + (\bar{w}_2 x)p_2 + (\bar{w}_2 y)q_2 + (\bar{w}_2)r_2 \qquad (11)$$

Which is a linear combination of the variable resultant parameters p1; q1; r1; p2; q2 and r2. The least squares method can be used to classify the optimal values of these parameters easily.

3.3 Classification using HNFRK

Subjects are classified into healthy control and AD using Hybrid Neuro Fuzzy Runge-Kutta (HNFRK) classifier. The ANFIS is a fuzzy sugenomodel placed in the adaptive system framework in order to facilitate adaptation and learning [23]. The ANFIS learns the feature vectors from the data set and the system parameters are adjusted according to a given error condition and the classifier is trained with Runge Kutta learning algorithm.

Given the training vectors the classifier classifies into two classes AD and healthy controls and assigns labels (y ϵ R^n) for every observation such that AD is considered as a positive class and Healthy controls are considered as negative class. Given the unlabeled data classes, the output of the classifier is as given as:

$$y = \sum_{i=1}^{n} x_i$$

where, n is the number of observations. The output of the classifier is the linear function of the defuzzifier parameters. The parameters are adjusted using the fourth order Runge Kutta learning algorithm [24]. The update mechanism relies on error back propagation. The neural networks based classification approach can be determined by the following equations:

$$\hat{x} = f(x, \tau)$$

$$x(t+1) = x(i) + \frac{h}{6}(p_0 + 2p_1 + 2p_2 + p_3)$$

$$p_0 = N(x; \propto) = N(x_0; \propto)$$

$$p_1 = N\left(x + \frac{1}{2}hp_0; \propto\right) = N(x_1; \propto)$$

$$p_2 = N\left(x + \frac{1}{2}hp_1; \propto\right) = N(x_2; \propto)$$

$$p_3 = N(x + hp_2; \propto) = N(x_3; \propto)$$

where, α is the generic parameter. Two paths are considered in this propagation. First the direct connection to the output summation and second via the neural network stages. Hence, each derivation, except the first derivation contains two terms. The fourth order Runge-Kutta approximation is summarized as follows:

$$\frac{\partial p_1}{\partial \propto} = \frac{\partial p_1}{\partial x_1}\frac{\partial x_1}{\partial p_0}\frac{\partial p_0}{\partial \propto} + \frac{\partial p_1}{\partial \propto}$$

$$\frac{\partial p_2}{\partial \propto} = \frac{\partial p_2}{\partial x_2}\frac{\partial x_2}{\partial p_1}\frac{\partial p_1}{\partial \propto} + \frac{\partial p_2}{\partial \propto}$$

$$\frac{\partial p_3}{\partial \propto} = \frac{\partial p_3}{\partial x_3}\frac{\partial x_3}{\partial p_2}\frac{\partial p_2}{\partial \propto} + \frac{\partial p_3}{\partial \propto}$$

$$\Delta_{\propto(i)} = \frac{rh}{6}(v^T(i) - x^T i))$$

$$(\partial p_0 + \frac{2\partial p_1}{\partial \propto} + \frac{2\partial p_2}{\partial \propto} + \frac{2\partial p_3}{\partial \propto})$$

Where r=The learning rate $v^T(i)$ = The measured state vector at the time t.

4 Result and Discussion

Three datasets are used here to evaluate the performance of the SSPCA gene selection algorithm. Two datasets were obtained from Keng Ridge BioMedical (http://datam.i2r.a-star.edu.sg/datasets/krbd) with one from the Tartu University Hospital. The genes are pre-filtered to reduce the number of genes to limit the computation time of the learning step and avoid singularities of the sample covariance matrix. Here, the pre-selection step uses the t-statistic filter on the training samples to retain 200 genes. The datasets are Leukemia, AML-ALL, Prostate cancer, Lung cancer.

A. Leukemia, AML-ALL

The data includes Acute Lymphoblastic Leukemia (ALL) and Acute Myeloid Leukemia (AML)[16]. Each sample was obtained from bone marrow analyzed using affymetrix microarrays containing 7129 genes. The training data consisted of 38 samples (1-27 are ALL and 28-38 are AML) while the test data had 34 samples (1-20 are ALL and 21-34 are AML).

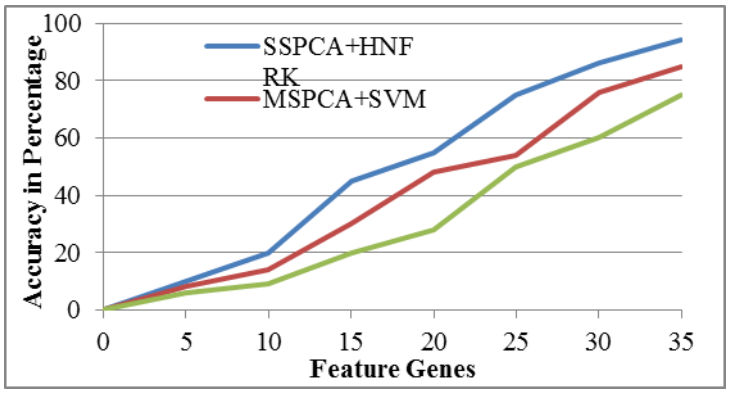

Figure 3: Comparison of accuracy using various approaches for Leukemia, AML-ALL dataset

Prostate cancer

The prostate cancer data consisted of 102 tissue training samples (1-52 prostate tumor and 53-102 normal tissues) and 34 tissue test samples (1-25 prostate tumor and 26-34 normal tissues) with 12 600 genes[24].

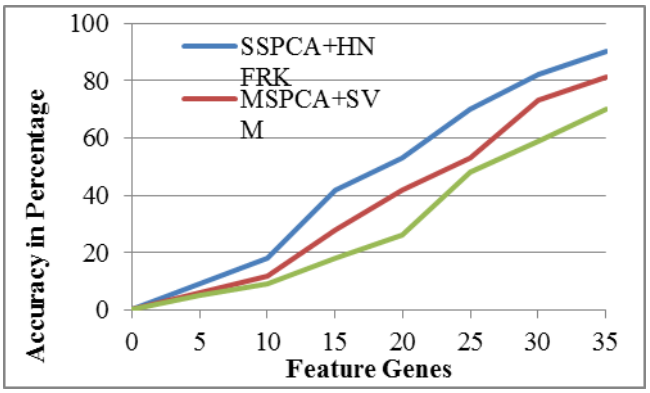

Figure 3: Comparison of accuracy using various approaches for prostate cancer dataset

B. Lung cancer

Lung cancer is one of the most common cancers worldwide and the leading contributor to cancer deaths, having one of the lowest survival rates within five years after diagnosis. The lung cancer data sample consisted of 18 072 genes and 106 samples from patients in Tartu University Hospital from November 2002 to December 2006, including 85 tumor samples (62 samples are squamous cell carcinoma and 23 samples are adenocarcinoma and its subtype bronchiole alveolar carcinoma) and 21 normal samples

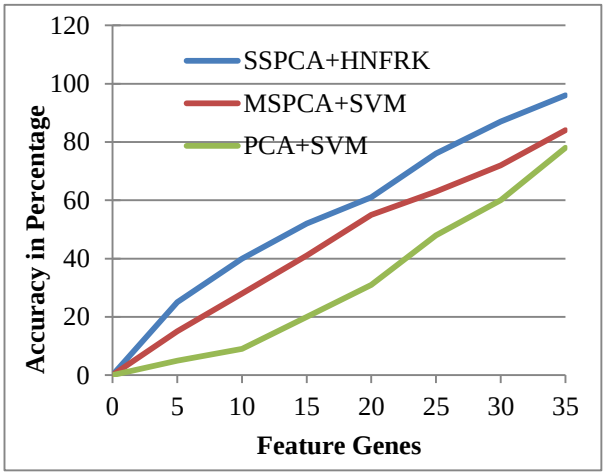

Figure 4: Comparison of accuracy using various approaches for Lung cancer dataset

5 Conclusions

Gene expression profiling by microarray technique has been effectively utilized for classification and diagnostic guessing of cancer nodules. Several machine learning and data mining techniques are presently applied for identifying cancer using gene expression data. In this paper, An SSPCA gene selection method was developed to select genes for classification. SSPCA was used to fit normal and cancer samples, with each element of each principal component loading used to select important genes by a sparse solution. Then the selection of genes undergoes the classification process by means of ELM classifier. An experimental result shows that, the proposed techniques like Structured Sparse Principal Component Analysis and the HNFRK Classifier gives better results than the other existing techniques.

References

C. Lazar, J. Taminau, S. Meganck, D. Steenhoff, A. Coletta, C. Molter, V. de Schaetzen, R. Duque, H. Bersini, and A. Now, A survey on filter techniques for feature selection in gene expression microarray analysis, Ieeeacm Trans. Comput. Biol. Bioinforma. Tcbb, vol. 9, no. 4, pp. 1106 1119, 2012.

A. El Akadi, A. Amine, A. El Ouardighi, and D. Aboutajdine, A twostage gene selection scheme utilizing MRMR filter and GA wrapper, Knowl. Inf. Syst., vol. 26, no. 3, pp. 487500, 2011.

T. Howlader and Y. P. Chaubey, Noise reduction of cDNA microarray images using complex wavelets, Image Process. Ieee Trans., vol. 19, no. 8, pp. 19531967, 2010.

N. Giannakeas, D. I. Fotiadis, and A. S. Politou, An automated method for gridding in microarray images, in Engineering in Medicine and Biology Society, 2006. EMBS06. 28th Annual International Conference of the IEEE, 2006, pp. 58765879.

L. Ying and C. Li, Based adaptive wavelet hidden Markov tree for microarray image enhancement, in Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. The 2nd International Conference on, 2008, pp.314317.

A. Figueroa, P. S. Tsai, E. Bent, and R. Guo, Robust spots finding in microarray images with distortions, in Engineering in Medicine and Biology Society, 2008EMBS 2008. 30th Annual International Conference of the IEEE, 2008, pp. 13391342.

Mohd Saberi Mohamad, Sigeru Omatu, Michifumi Yoshioka and Safaai Deris, "A THREE-STAGE METHOD TO SELECT INFORMATIVE GENES FOR CANCER CLASSIFICATION", International Journal of Innovative Computing, Information and Control ICIC International c2010 ISSN 1349-4198 Volume 6, Number 1, January 2010.

I.T. Jolliffe,Principal Component Analysis.2nd edition. New York:Springer,2002.

Rodolphe Jenatton, Guillaume Obozinski, Francis Bach, "Structured Sparse Principal Component Analysis", hal-00414158, version 3 - 8 Sep 2009.

J.B.Tenenbaum, V.Silva and J.C.Langford, A global geometric framework for nonlinear dimensionality reduction,vol.290,2000,pp.2319-2323.

S.T.Roweis,L.K.Saul,Nonlinear dimensionality reduction by locally linear embedding,vol.290,2000,pp.2323-2326.

X. He and P. Niyogi, Locality Preserving Projections, Advances in Neural Information Processing Systems 16, , Cambridge,MIT Press, 2003.

C. Shi and L.H. Chen, Feature Dimension Reduction for Microarray Data Analysis using Locally Linear Embedding. APBC,vol. 16, 2004,pp.1-7.

G. Lee, C. Rodriguez and A. Madabhushi , Investigating the Efficacy of Nonlinear Dimensionality Reduction Schemes in Classifyi Gene- and Protein-Expression Studies, IEEE/ACM Transactions on Computational Biology and Bioinformatics,vol.5,2008, pp. -384.

K. Huang and S. Aviyente, Sparse Representation for Signal Classification, Advances in Neural Information Processing Systems, vol.19,2006, pp. 609-616.

S. Yan and H. Wang, Semi-Supervised Learning by Sparse Representation, SIAM International Conference on Data Mining, pp. 792-801 March,2009.

Chunming Xu, "Gene Expression Data Classification Using Discriminatively Regularized Sparse Subspace Learning", World Academy of Science, Engineering and Technology 51, 2011.

Ronny Luss, Alexandre d'Aspremont, "Clustering and Feature Selection using Sparse Principal Component Analysis", July 4, 2007.

Yanwei Huang, Liqing Zhang, "Gene Selection for Classifications Using Multiple PCA with Sparsity", TSINGHUA SCIENCE AND TECHNOLOGY, ISSNll1007-0214ll06/10llpp659-665, Volume 17, Number 6, December 2012.

R. Jenatton, J-Y. Audibert, and F. Bach. Structured variable selection with sparsity-inducing norms. Technical report, 2009. preprint arXiv:0904.3523v1.

Wang A, Gehan E A. Gene selection for microarray data analysis using principal component analysis. Statistics in Medicine, 2005, 24(13): 2069-2087.

Jang, J-SR. "ANFIS: adaptive-network-based fuzzy inference system."Systems, Man and Cybernetics, IEEE Transactions on 23.3 (1993): 665-685.

Boyacioglu, M.A. and D. Avci, 2010. An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of the Istanbul stock exchange. Expert Syst. Appl., 37(12): 7908-7912.

Musa, H., I. Saidu and M.Y. Wazir, 2010. A simplified derivation and analysis of fourth order runge kutta method. Int. J. Comput. Appl., 9(8): 51-55.