Calculating the Cancer Gene Expression Data Using Radial Basis Function and Neural Network: A Survey

R. Dhivya[1], G. Sumalatha[2], M. Swathinayaki[3], P. Sharon silviya[4]

[1,3,4]Student, Dept. of Computer Technology, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India.

[2] Assistant Professor, Dept. of Computer Technology, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India.

Abstract

Cancer is a very much major problem which gives tough to the medical science throughout in the world. Now a days so many data mining techniques are used to detect the Cancer in their earlier stages are increasing. But getting the data accuracy gives the very poor results in researching to know the idea about the Cancer, accurate classification is more essential for successful diagnosis of Cancer and helps to the doctors to start the proper treatment. Cancer gives the cell mutation in genetic structure. So much uncontrolled growth of these patterns and cells are difficult to identify. Gene expression gives the information's about the regulated genetic levels and to monitor the thousands of genes simultaneously. This survey gives the details about calculating the gene expression to detect the cancer in their starting point by the Radial Basis Function and Neural Network. In future modifications of this technique can produce 100% accuracy and proper methods to cure cancer.

Key Words :_Gene mutation, gene network and gene expression, radius basis function, neural network._

I. INTRODUCTION

In this advanced world we can find the cause of every disease and genome problems. Gene expression gives the information's about major problems like cancer. The accuracy of the techniques are very much use full to predict the disease in their early stage. Gene information are gives the particular function elements such as gene product. Large amount abnormal of these products will cause the oncogene. In this formation the nucleic acids(nucleotides) sequences in DNA or RNA are altered. Gene mutations are block any particular are of a DNA. Data sets of these mutated genes are processed with a algorithms. High level of accuracy given by the radius basis algorithms and neural networks in the gene expression network and data. The review of accuracy values for the gene expression are explained in this paper.

In the process of meiosis, the errors of DNA replications and the metabolic activity, radiations, TLS are affects the gene. Gene expression gives the lot of query genes simultaneously. But in such cases we cannot get the combine genes query values. It gives only damaged and affected genes queries. The

patient's clinical outcomes are predicted by the derived information's. An organism of all the cells contain the organism of the entire genome , and the small subset of the genes are expressed with messenger RNA (mRNA) at any time, and the relative expression of them can be evaluated.

Messenger RNA (mRNA) contains lot of RNA molecules. In is the genetic information from the DNA to the next ribosome are conveyed. They specified the sequence of amino acid in the protein products of gene expression . In DNA, mRNA the nucleotides are arranged into codons .Protein synthesis are terminated by the stop codens. This process has two types Transfer RNA (tRNA), the corresponding amino acid are provided by them. Ribosomal RNA (rRNA), that is the protein-manufacturing machinery. mRNA process the enzymes it is happened quickly on the start of transcription.

DNA microarrays are collection of a DNA microscopic spot which is used to measure the level of expression in simultaneously. Data sets of microarrays ray are very large, and a number of variables are influenced the analytical precision and they provide the normalization of the data without any noise .The methods may be suited to specific platforms only.The number of genes which is used by the classifier and the classification accuracy are used to evaluate the cancer classifications. The accuracy of the classifications are increased by the number of genes which are reduced or irrelevant. This reduced genes are the noise for every classifier. Relevent of small genes are used to found the development of the cancer and the practical importance.

The classification of these gene expression data are more power and compulsory processes to identify the large amount of abnormal genes. Computer chips are used to seen the digit network of activation functions and they are ON or OFF depending on their input. Time series analysis gives the explicit instructions, patterns relating unsupervised learning to the group, or segment, datasets with shared attributes in that which is used to order to extrapolate the algorithmic relationships. And also the accuracy level reached 100% by the Radial techniques so that the doctor can easyly detect the early stage of the cancer and we can predict them in thir starting stage. The time series analysis are used to get the quick recovery stages of the cancer.

II. LITERATURE SURVEY:

Bala Chandra, KV NareshBabu produce A new spiking function has been proposed in the non-linear integrate and fire model and it's inter spike interval is derived and used in the Wavelet Radial Basis Neural Network for the classification of gene expression data.They gives the comparative performance between the WRNN and SWRNN. The Spiking function gives the powerfull classification.

Feng Chu and Lipo Wang proposed RBF Neural network for lymphoma, SRBCT and ovarian cancer data sets and they produces 100% accuracy for lymphoma with 9 genes, SRBCT with 8 genes and ovarian cancer with 4 genes. They used two techniques. First they collect the training data set and then they apply the data sets with the Radial Basis Function. MaciejKusy, BogdanObrzut and JacekKluska used 47 cancer cervical patients data sets and compares the gene expression programming(GEP)and the multilayer perceptron (MLP). They produce the 71.96% accuracy for gene expression programming for the MLP 71.87% accuracy. M.VenkateswaraRao,N.V.RamanaMurty proposed the ANN and RBF neural networks for lung cancer and they collects 50 informations 25 for cancerous and 25 for non cancerous. The proposed Radial Basis Function Neural Network gives the 97% result with the highest TP Rate it used to identify the lung cancer in early stage.

J. Padmaavati presented the comparative prediction level for the cancer and they used 580 records and they gives correct prediction for MLP with 91.3%and the RBF with 97.0%. They gives the Sensitivity for MLP is 92.1% and RBF with 97.3 for

the specificity MLP with 91.% and RBF with 96.8%.

John shell, William D.Gregory presents the multiple neural networks for the cancer detection and they takes the data sets of 180 malignant, 180 benign breast tissue. The propesed BNN gives the 100% sensitivity and the specificity the variations between the two data sets are configured by the neural networks.

Emmanuel Adetiba and Oludayo O. Olugbaracompares the SVM and ANN for lung cancer prediction. They collected the p53 genome samples. They apply the HOG and LBP to the encoded nucleotides. The ANN with HOG produces the 95.90% accuracy with the smallest mean square error of 0.0159.

Jun S. Wei, Braden T. Greer and Javed Khan are use the ANN for predict the neuroblastoma. In this 37920 microarray clones gives the 88% accuracy. The predicted accuracy 98% is given by only 19 genes.

Yen-Chen Chen, Yo-Cheng Chang presents the ACT correlated with ANN for non small cell lung cancer and they gives the 65.71% accuracy with the 2 genes. They applies the ACT methods for used to prevent the metastases after the surgery.

Sung-BaeCho,Hong-Hee Won used the neural network method for predict leukemia with 95.9% accuracy in 38 samples out of 72 and they use the 40 samples for colon cancer and produce the results for 87.9%. These are used to predict them in the starting stage.

CONCLUSION

Cancer creates the causes of quick dissemination of the cells and gene. The high dimensionality of gene expression data of microarray creates the method needs and automatically detect the structures in the data. In the Radial Basis time series analysis used to predict the causes of the cancer in the initial condition. The normalized critical stage of this mutation are represented and the doctor can easily know the development of the cancer genes.

There are so many machine learning techniques are used to find the DNA microarrays and the gene profiling. Neural network gives the good results and gives the correct sensitivity to identify the expression and the gene current situation. In the case of prediction we can easily calculate the expressions and the three parts such as sensitivity, specificity and the prediction levels by using the RBF-NN. In future research of the modern techniques which contains the RBF will produce easy detection and control systems. And gives the correct accuracy level and also give the proper treatment after the surgery also.

REFERENCE :

[1] A. B. Olshen, and A. N. Jain, "Deriving quantitative conclusions from microarray expression data,"2002.

[2] Microarray Technology and Cancer Gene Profiling published on 2007.

[3] DNA Microarrays:Gene Expression Applications published on 2001

[4] Advanced Analysis of Gene Expression Microarray Data Book by Aidong Zhang published on 2006.

[5] Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, et al."Complementary DNA sequencing: expressed sequence tags and human genome project"(June 1991).