Abstract

Recent advances in the DNA microarray technology have provided the ability to examine and measure the expression levels of thousands of genes simultaneously in an organism. In this technology each gene is recorded under different conditions or each gene is evaluated under a single environment but in different types of tissues. In the first case it is used in identification of functionally related genes whereas the second type of technology is helpful in classification of different types of tissues and identification of those genes whose expression levels are good diagnostic indicators. Different approaches have been applied to classify different datasets. However, the main challenges in this task is the availability of a smaller number of samples compared to huge number of genes and the noisy nature of biological data. This paper reviews different techniques used to classify the genes and improved efficiency of biomarker identification due to these classifications.

References

Review on Classification of Genes and Biomarker Identification

- Feng Chu and Lipo Wang, "Applying Rbf Neural Networks To Cancer Classification Based On Gene Expressions," International Joint Conference on Neural Networks, July 16-21, 2006.
Review on Classification of Genes and Biomarker Identification

- Jinn-Yi Yeh, Tai-Shi Wu, Min-Che Wu and Der-Ming Chang, "Applying Data Mining Techniques for Cancer Classification from Gene Expression Data", International Conference on Convergence Information Technology, 2007.
- M. Maruf Hossain, Md. Rafiul Hassan and James Bailey, "A Novel Decision Tree Induction Algorithm Based on Receiver Operating Characteristics to Classify Gene Expression Data", The University of Melbourne, Australia.
- Oscar Picchi Netto, Sergio Ricardo Nozawa Rafael Andrade Rosales Mitrowsky, Alessandra Alaniz Macedo 1, Jose Augusto Baranauskas, "Applying Decision Trees to Gene Expression Data from DNA Microarrays: A Leukemia Case Study", 20-23 July 2010, ISSN 2175-2761.

Index Terms

Computer Science   Bio-informatics

Keywords

Data mining  DNA microarray  Support vector machine (SVM)  Decision tree
Neural network

Biomarkers