Open access Journal International Journal of Emerging Trends in Science and Technology

Supervised Multi Attribute Gene Manipulation for Cancer

Authors
R. Mahalakshmi[1], I. Mohanapriya[2], K. Kasiyammal[3], Ms. V.Sathiya M.E. [4]
Panimalar Engineering College, Chennai, Tamil Nadu
Email: *mahamona2694@gmail.com[1],priyailango1989@gmail.com[2], chitrakasi24@gmail.com[3], deviviji2000@yahoo.co.in[4]*

ABSTRACT
Cancer research is the one of the major research areas in the medical field. Pointed out the exact tumour types provides an optimized solution for the better treatment and toxicity minimization due to medicines on the patients. To get a clear picture on the insight of a problem, a clear cancer classification analysis system needs to be pictured followed by a systematic approach to analyse Global Gene Expression which provides an optimized solution for the identified problem area. Molecular diagnostics provides a promising option of systematic human cancer classification, but these tests are not widely applied because characteristic molecular markers for most solid tumours have yet to be identified. Recently, DNA microarray-based tumour gene expression profiles have been used for cancer diagnosis. Existing system focussed in ranging from old nearest neighbour analysis to support vector machine manipulation for the learning portion of the classification model. Supervised Multi Attribute Clustering Algorithm, which can manage knowledge, attributes coming two different knowledge streams. Our proposed system takes the input from multiple sources, creates an ontological store, cluster the data with attribute match association rule and followed by classification with the knowledge acquired.

Key Words: *Global Gene Expression, Supervised Multi Attribute Clustering Algorithm, Ontological store.*

INTRODUCTION

Data mining is one of the major research areas in the field of medical. It is a process to extract the implicit information and knowledge by extracting from the mass, incomplete, noisy, fuzzy and random data with knowing the data well in advance and which is potentially useful to various fields. The goal of data mining is to extract knowledge from dataset in human-understandable structures.

In Health care sector, it supports by correlating patients with critical illnesses, developing better insights on symptoms and their causes and learning how to provide proper treatments. CANCER is a major cause of all the natural mortalities and morbidities throughout the world. Nearly 13 percent of deaths caused are due to cancer. It is a disease getting constantly challenged by many eminent and premier researchers. The growth in a body is observed when the division and multiplication of cells takes place. When the appropriate division levels have been achieved, the process is deactivated. In an unusual scenario however, cells continue to replicate and form lumps in the body, although it commences with a paltry entity.

Cancer is an abnormal and uncontrollable growth of cells in the body that turn malignant. This is not to be confused with tumors.

Even a tumor is an abnormal growth of cells, but it can be classified as (noncancerous) benign and malignant, the latter one causing cancer. Various types of cancer have been identified namely, breast cancer, colon cancer, lung cancer, brain cancer, cervical cancer, kidney cancer, liver cancer, leukemia, Hodgkin's lymphoma, non-

Hodgkin's lymphoma, ovarian cancer, skin cancer, thyroid cancer, uterine cancer, and testicular cancer. Cancer causes quick dissemination of cells and a cancer type can fortify and extend to another one if not treated appropriately. Symptoms of cancer depend on the type and location of the cancer. For example, lung cancer can cause coughing, heavy breathing, chest pain, etc.

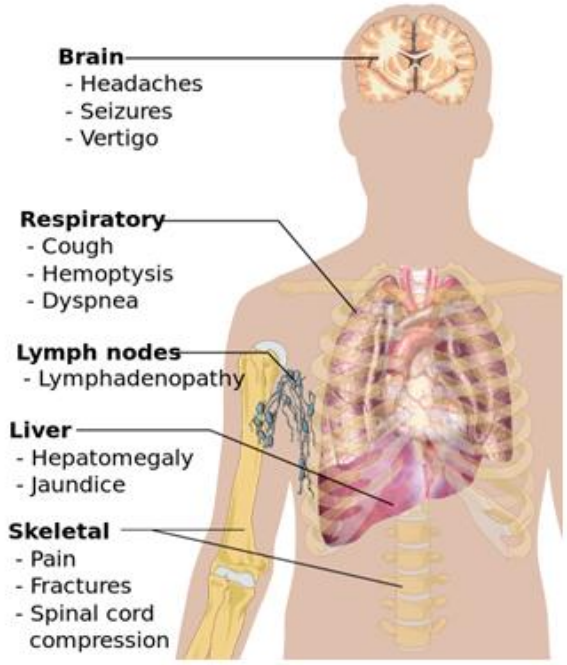

Figure 1: Cancer Metastasis

Some cancers may not have any symptoms at all. In certain cancers, such as pancreatic cancer, symptoms often do not start until the disease has reached an advanced stage. Cancer progression can be aggressive or benign. That corresponds to the suitable treatment required for ailing the cancer.

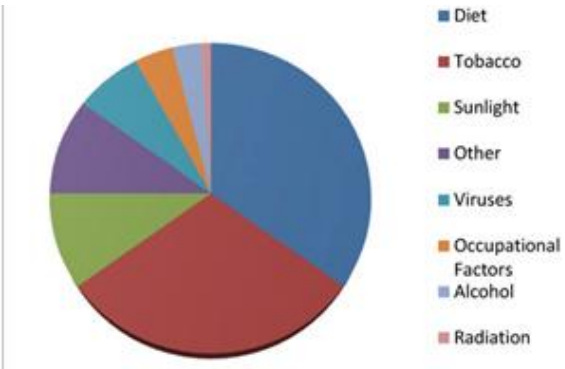

Figure 2: Causes of Cancer

RELATED WORK

Gene expression is the activation of a gene that results in a protein which tends to identify only the Gene manipulation for cancer therapeutics. In our existing approach, identification of cancer by the gene expression has been implemented. The Genome of a Differentiated Cell Contains all the Genes required finding the affected cells using Microarrays to investigate the "Expression" of Thousands of Genes at a Time. 1. Splicing, 2.Polyadenylation, 3.Stability. Discretized gene expressions can be used as descriptors of the specific states of gene for the cancer prediction analysis. This disadvantage of this method is Fragments of transposons are often confused for protein-coding exons of genes. Direct DNA Injections leads to Low Efficiency identification

PROPOSED WORK

There are so many data mining techniques existing today from which we have proposed our method. In this method system takes the input from multiple supply, produce associate metaphysics store, cluster the info with attribute match association rule and followed by classification with the information non inheritable. To get a transparent image on the sight of a tangle, a transparent cancer classification associate analysis system has to be pictured followed by a scientific approach to analyze international organic phenomenon that provides an optimized answer for the known drawback space. Molecular medical specialty provides a promising possibility of systematic human cancer classification, however these tests don't seem to be wide applied as a result of characteristic molecular markers for many solid neoplasm save nevertheless to be known. Predicting Cancer by analyzing gene and converting the gene expression is the proposed concept of our project, which leads to identifying and analyzing the cancer result set.

Controlling Gene Activity from Gene to Functional Protein & Phenotype has also been analyzed in order to identify the cancer cells. In our proposed methodology, the expert's documental DNA data methylation (Gene

expression segments) is a kind of binding site for proteins which make DNA inaccessible to be in alive state. Semantic Ontology based Mining Gene Expression analysis tends to compare the gene expression values by using the comparative Knowledge Consolidator. Supervised Multi Attribute Clustering Algorithm has been used to find the Best Rule Classification in the gene expression to find the Final Prediction of cancer disease.

The advantage of this method is Comparing genomes can help with gene finding. Can help visualize transient gene expression and to identify if tissue is stably transgenic and useful for cellular and ecological studies.

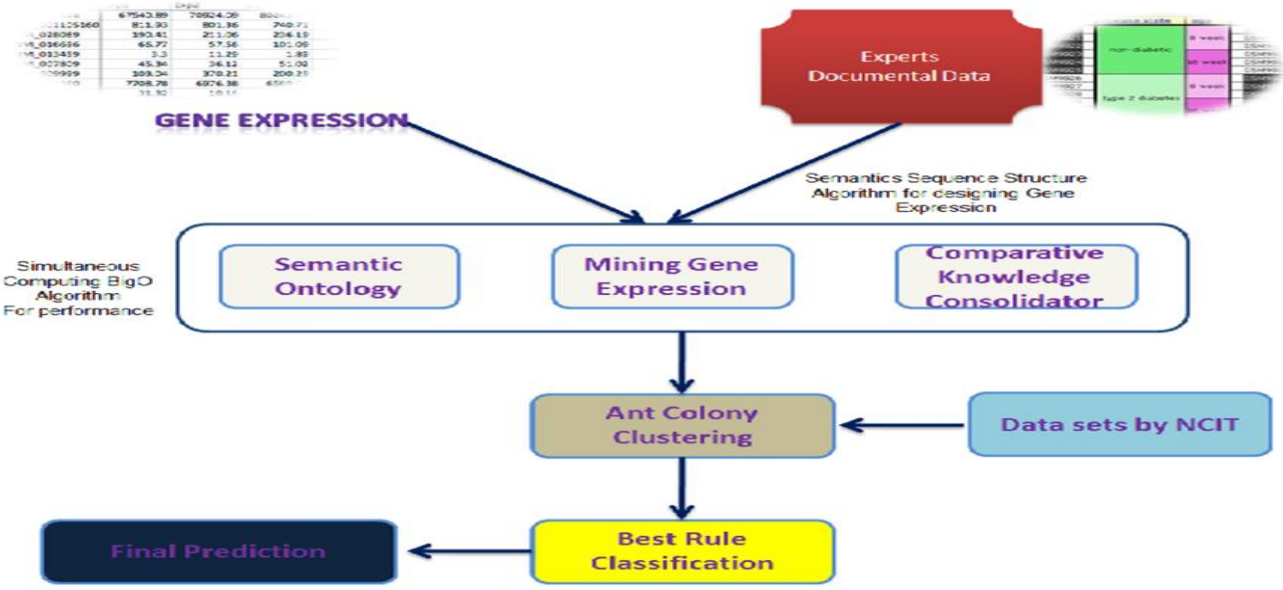

Figure 3: Architecture diagram for Mining Gene Expression Cancer Therapeutics

SAMPLE SCREENSHOTS

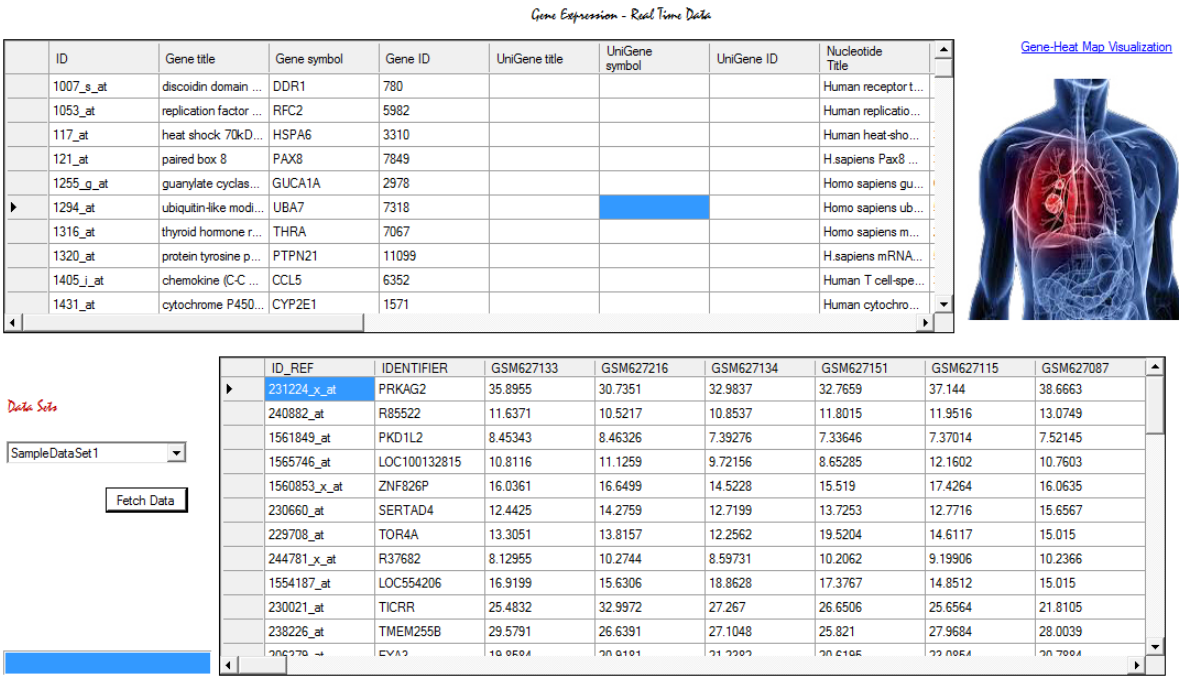

Figure 4: Select Dataset And Fetch Data

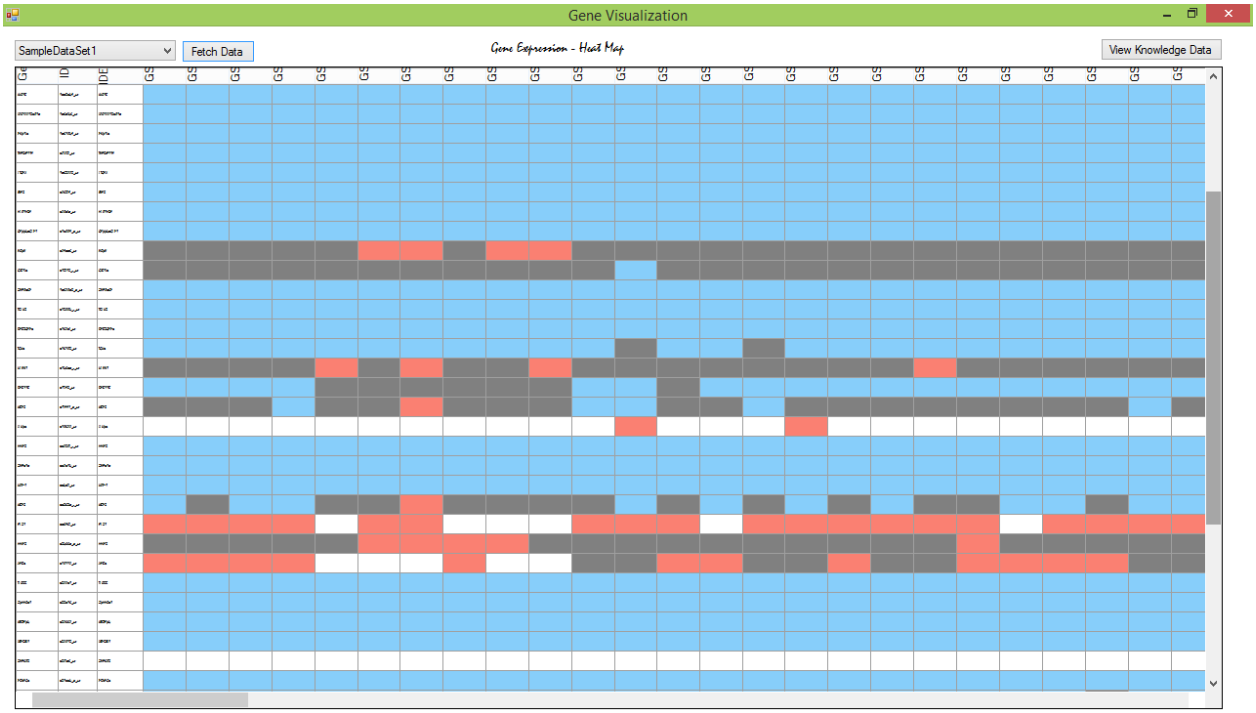

Figure 5: Gene Heat Map Visualization

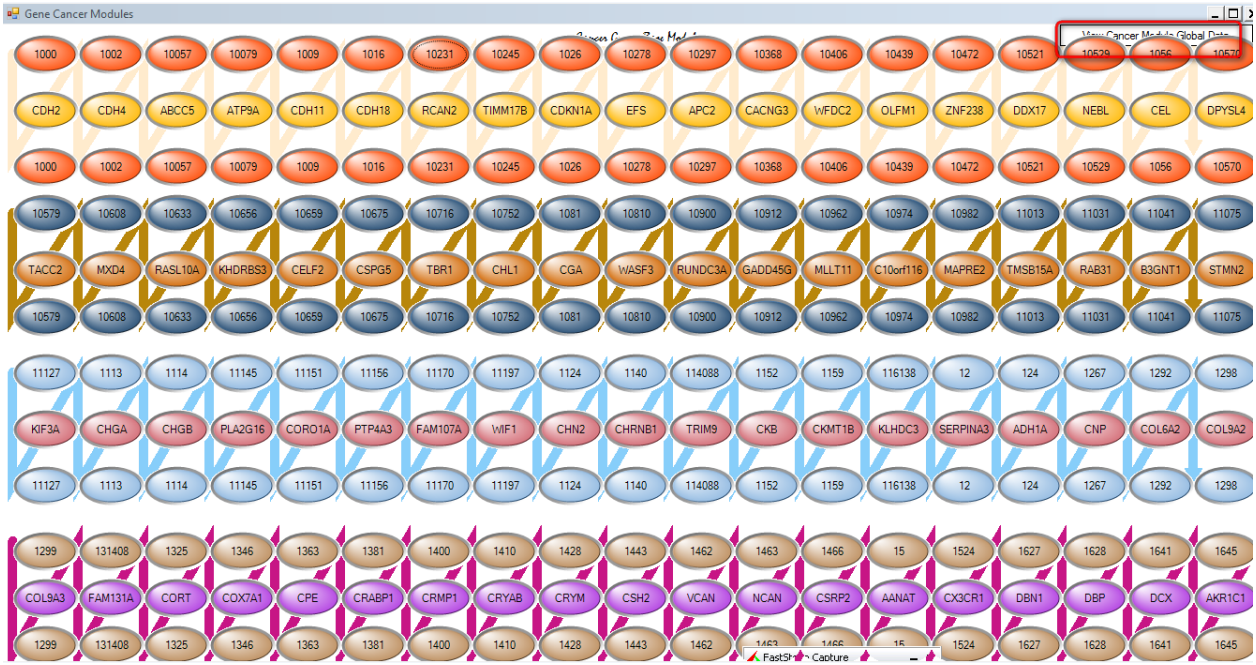

Figure 6: Dataset Patterns

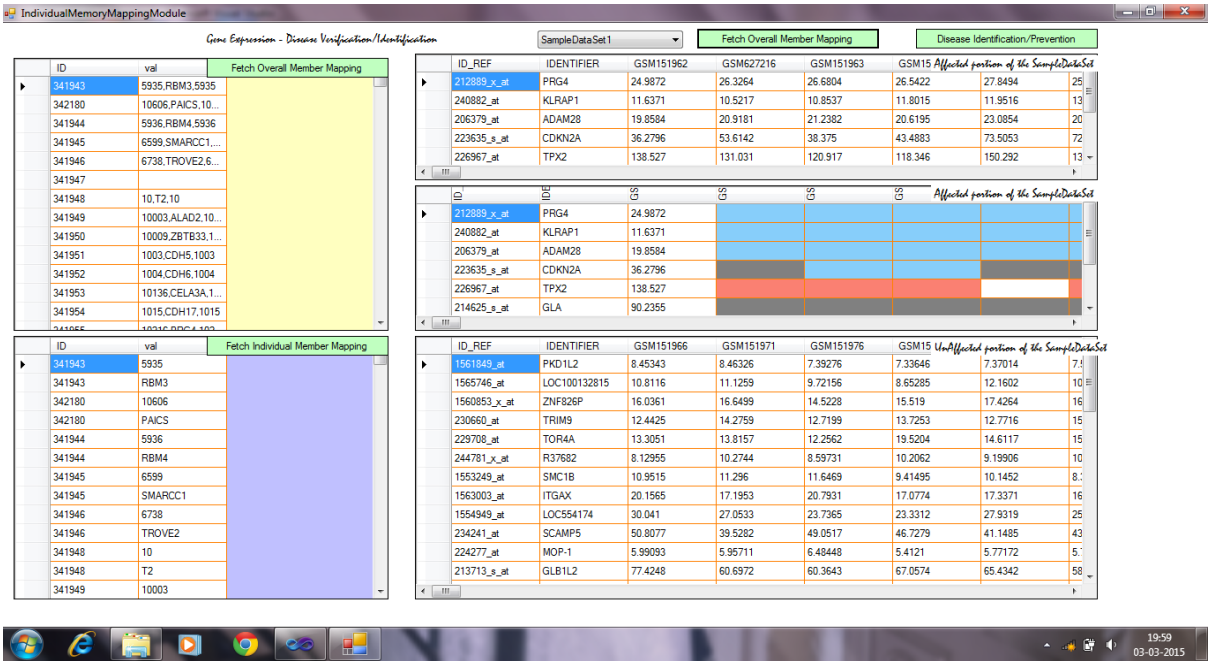

Figure 7: Identification of affected and unaffected portion

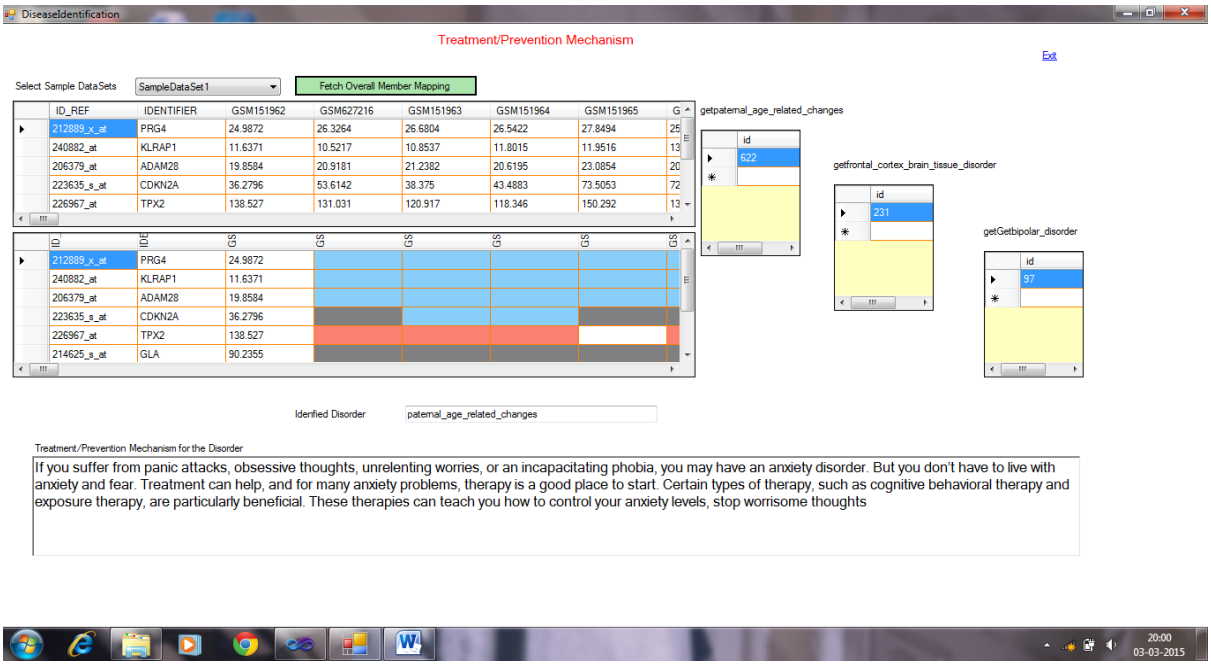

Figure 8: Disorder identification (Secondary Cancer)

CONCLUSION

This paper will be produce the classification of tumors is essential for successfully identify the diseases and treatment of cancer And monitoring of expression levels in cells for thousands of genes simultaneously and Also annotated is that comparing the activity of genes in a healthy and cancerous tissue genes that are involved in cancer. And also find the deep of affected genes and also give the solution for the cancer affected genes.

FUTURE ENHANCEMENT

Cloning, body parts, or even entire bodies generally referred to as reproductive cloning that would be genetically identical to a prospective patient. Recently, the US Department of Defense initiated a program to research the possibility of growing human body parts on mice. Complex biological structures, such as mammalian joints and limbs, have not yet been replicated. The implantation of bio-engineered bladders grown from patients' own cells has proven to be a viable

treatment for bladder disease. Proponents of body part replacement and cloning contend that the required biotechnologies are likely to appear earlier than other life-extension technologies.

REFERENCES

1. Data and Statistics. World Health Organization, Geneva, Switzerland, 2006.
2. PubMedHealth- U.S. Nat. Library Med., (2012). [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0002267/
3. S. Dudoit, J. Fridlyand, and T. P. Speed, "Comparison of discrimination methods for the classification of tumors using gene expression data," J. Amer. Statist. Assoc., vol. 97, no. 457, pp. 77–87, Mar. 2002.
4. G.-M. Elizabeth and P. Giovanni, (2004, Dec.). "Clustering and classification methods for gene expression data analysis." Johns Hopkins Univ., Dept. of Biostatist. Working Papers. Working Paper 70. [Online]. Available: http://biostats.bepress. com/jhubiostat/paper70.
5. E. Shay, (2003, Jan.). "Microarray cluster analysis and applications" [Online]. Available: http://www.science.co.il/enuka/ Essays/ Microarray-Review.pdf.
6. M. B. Eisen, T. P. Spellman, P. O. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," Proc. Nat. Acad. Sci. USA, vol. 95, no. 25, pp. 14863–14868, Dec. 1998.
7. S. Tavazoie, D. Hughes, M. J. Campbell, R. J. Cho, and G. M. Church, "Systematic determination of genetic network architecture," Nature Genetics, vol. 22, pp. 281–285, 1999.
8. T. Kohonen, Self-Organising Maps. Berlin, Germany: Springer- Verlag, 1995.
9. N. Pasquier, C. Pasquier, L. Brisson, and M. Collard, (2008). "Mining gene expression data using domain knowledge," Int. J. Softw. Informat, vol. 2, no. 2, pp. 215–231, [Online] Available: http://www.ijsi.org/1673-7288/2/215.pdf.
10. N. Revathy and R. Amalraj, "Accurate cancer classification using expressions of very few genes," Int. J. Comput. Appl., vol. 14, no. 4, pp. 19–22, Jan. 2011.