

## International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology



ISSN 2320-088X

International Conference on Mobility in Computing- ICMiC13, Organized by Mar Baselios College of Engineering and Technology during December 17-18, 2013 at Trivandrum, Kerala, India, pg.180 – 190

### **REVIEW ARTICLE**

# Review of Cancer Subtypes Prediction: Progress and Challenges

Gopika Rani G. U.<sup>1</sup>, J. Jesu Vedha Nayahi<sup>2</sup>

<sup>1,2</sup>Computer Science and Engineering, Regional Centre of Anna University, Tirunelveli, Tamilnadu, India

<sup>1</sup>gugopikarani@gmail.com

---

*Abstract— Microarray cancer data, organized as samples versus genes fashion, are being exploited for the classification of tissue samples into benign and malignant or their subtypes. They are also useful for identifying potential gene markers for each cancer subtype, which helps in successful diagnosis of particular cancer type. Nevertheless, small sample size remains a bottleneck to design suitable classifiers. Traditional supervised classifiers can only work with labelled data. On the other hand, a large number of microarray data that do not have adequate follow-up information are disregarded. This survey paper categorises, compares, and summarises the various approaches used in the field of cancer subtypes prediction in Gene Expression based Biological data mining. It defines the techniques used for cancer subtype prediction that includes the basic techniques like decision trees, rules, etc to the recent methodologies based on SVM. The association of feature selection methods with the classification methods is also included in this study which gives better performance in terms of accuracy and time complexity. Recent studies on cancer prediction introduce the scope of semi-supervised learning. This study presents methods and techniques in the field of semi-supervised cancer subtypes prediction incorporated with feature selection mechanisms and their issues. Compared to all related reviews, this survey proposes a novel approach to combine feature (gene) selection and transductive support vector machine (TSVM) where potential gene markers can be identified and TSVMs can improve prediction accuracy as compared to the standard inductive SVMs (ISVMs).*

*Keywords— Gene expression profiling; semi-supervised learning; gene selection; consistency; TSVM*

---

Full Text: <http://www.ijcsmc.com/docs/papers/ICMIC13/ICMIC13S21.pdf>