

## A Review of Cancer Classification Software for Gene Expression Data

Tan Ching Siang<sup>1</sup>, Ting Wai Soon<sup>a</sup>, Shahreen Kasim<sup>b</sup>, Mohd Saberi Mohamad<sup>a</sup>, Chan Weng Howe<sup>a</sup>, Safaai Deris<sup>a</sup>, Zalmiyah Zakaria<sup>a</sup>, Zuraini Ali Shah<sup>a</sup> and Zuwairie Ibrahim<sup>c</sup>

<sup>a</sup>Artificial Intelligence and Bioinformatics Research Group, Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor.

<sup>b</sup>Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor

<sup>c</sup>Faculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang.

### ABSTRACT

Microarray technology provides a way for researchers to measure the expression level of thousands of genes simultaneously in a single experiment. Due to the increasing amount of microarray data, the field of microarray data analysis has become a major topic among researchers. One of the examples of microarray data analysis is classification. Classification is the process of determining the classes for samples. The goal of classification is to identify the differentially expressed genes so that these genes can be used to predict the classes for new samples. In order to perform the tasks of classification of microarray data, classification software is required for effective classification and analysis of large-scale data. This paper reviews numerous classification software applications for gene expression data. In this paper, the reviewed software can be categorized into six supervised classification methods: Support Vector Machine, K-Nearest Neighbour, Neural Network, Linear Discriminant Analysis, Bayesian Classifier, and Random Forest.

**KEYWORDS:** Cancer Classification, Gene Expression Data, Microarray, Supervised Classification Methods, Bioinformatics, Artificial Intelligence