

ABSTRACT

Microarray is a modern technique facilitating the simulation analysis of a number of large gene expression data which is needed to solve many complex biological problems. Therefore, there is a scheme which required the process of dimensional reduction and classification process. In this case, the dimension reduction process aims to relieve the computational load the classification, the reduction process used is the feature selection of Genetic Algorithm and feature extraction of Wavelet Haar. The process of classification aims to classify data, whether it is cancer or not, by using the Naïve Bayes classification method. As a result, the accuracy of Genetic Algorithm selection of the data features of detection of Colon Tumor is 76,4706%, 98.0769% of Lung Cancer, and 75% for ovarian. While, Haar Wavelet feature extraction from Colon Tumor data has 80% of accuracy, 94,1176% of Lung Cancer and 100% for ovarian.

Keywords: *microarray data, naïve bayes, genetic algorithm, haar wavelet*