

ABSTRACT

Cancer is a disease that can cause human deaths in the world. Based on data obtained from GLOBOCAN, the International Agency for Research on Cancer (IARC) in 2012, it is known that there were 14.1 million new cases, 8.2 million cancer deaths [1]. Therefore, we need a DNA microarray technology which in recent years is often used as a technology to analyze and diagnose cancer. With the analysis of gene expression data in the form of the microarray, allowing a medical to be able to know whether a person has a cancer or not. Data in DNA microarray has huge dimensions, the magnitude of the dimensions can affect the cancer classification process. Therefore, we need a scheme which there is a dimension reduction process and classification process, so that the microarray data classification scheme to get results and good accuracy. Many methods can be applied in cancer classification based on microarray data, the one Support Vector Machines (SVM) as a method of classification and Principal Component Analysis (PCA) as a method of dimension reduction, both of which have been tested in several previous studies. By applying these methods, the method of PCA and SVM method on several types of cancer data, it was found that the cancer data central nervous, colon, ovarian, lung, breast and leukemia get results more than 80% accuracy for the linear kernel and RBF

Keywords: cancer, classification, dimension reduction, SVM, PCA.