

Abstract

Cancer is one of the leading causes of human death in the world. It is estimated that cancer patients continue to increase every year. Cancer that can be detected early has a higher probability of getting a faster and more appropriate treatment. One way is by using Microarray technology. Microarray technology can analyze gene expression profiles at the same time. By analyzing the Microarray data, it can then be known whether a person has cancer or not. However, the problem with Microarray data is that there are far more attributes than samples, so dimension reduction is necessary. To overcome this, the authors use one of the dimensional reduction techniques, namely Principal Component Analysis (PCA) and 2 classification methods, namely Naïve Bayes and Support Vector Machine (SVM), which will then be compared and analyzed the performance results of the two methods to find out which one is the best. better. The accuracy of the results of this study shows that 4 out of 5 cancer data get an accuracy of 77-96% while 1 other data, namely breast cancer data, gets the smallest value, which is 54.6%.

Keywords: Cancer, Microarray, Dimension Reduction, Principal Component Analysis (PCA), Naïve Bayes, Support Vector Machine (SVM).