

Abstract

According to WHO data in 2015 (World Health Organization), 8.8 million deaths were caused by cancer where the mortality rate increased and was fatal every year if the diagnosis was not made earlier. Therefore, it is not surprising that research in the field of cancer has become a major topic in research in the medical and bioinformatics fields and continues to grow to date, including DNA microarray technology. There are many ways to detect cancer, one of which is the microarray technique. Microarray is a technology that can store thousands of gene expressions taken from several human tissues at once. Due to a large number of microarray data records, the computing required is quite heavy. To overcome this problem, dimension reduction is needed. In this study, the system uses two extraction features: Principal Component Analysis (PCA) and Partial Least Square (PLS) with Support Vector Machine (SVM) as a classifier. This is useful to reduce the large amount of attributes. The accuracy generated from this system averaged 82% with PCA-SVM and 55.17% for PLS-SVM.

Keywords: cancer, microarray, principal component analysis, partial least square, support vector machine.