

Abstract

Cancer is the second largest cause of death in the world; in 2015, a total of 8.8 million mortalities were recorded, due to cancer alone. It is important to detect this deadly disease early. In the medical field, there are many methods that can be used to detect cancer. One of these methods is microarray data technology. Microarray data reads thousands of gene expressions at the same time. However, this method has a major problem; data with high dimensions can affect classification performance and consume a lot of computational time. Therefore, this research used Principal Component Analysis as the dimensional reduction method. This method performed feature extraction based on a Principal Component (PC) obtained from the calculation of eigenvalues and eigenvectors. After applying dimensional reduction, the data was classified using a multinomial logit classifier. The cancer data used in this research consists of Colon Cancer, Leukemia, Lung Cancer, and Ovarian Cancer datasets. The test results for the Ovarian Cancer dataset gave an accuracy of 100% using a Proportion of Variance (PPV) of 90%.