Implementasi Partial Least Square dan K-Nearest Neighbor - Support Vector Machines Untuk Klasifikasi Data Microarray

A Rakha Ahmad Taufiq[1], Adiwijaya[2], Annisa Aditsania[3]

[1,2,3]Fakultas Informatika, Universitas Telkom, Bandung
[1]ahmadtaufiq@students.telkomuniversity.ac.id, [2]adiwijaya@telkomuniversity.ac.id,
[3]aaditsania@telkomuniversity.ac.id

Abstract
Cancer is one of the most common causes of death in the world. Estimated every year the number will continue to grow. One of the detection is using gene expression. Microarray can collect a large number of gene expression at the same time, DNA Microarray have their own data characteristic, which have a very large data dimension compared with the amount of data. Therefore, a system needed to solve the problem. In this research, we built a system that implements Partial Least Square (PLS) feature extraction and K-Nearest Neighbor - Support Vector Machines (KNN-SVM) for the classification. Feature extraction is useful for reducing very large dimension of microarray by forming new data. System performance is measured using accuracy. PLS managed to increase the accuracy of the KNN-SVM classifier. The highest accuracy obtained by PLS KNN-SVM is 96.17%.

Keywords: k-nearest neighbor, support vector machines, partial least square, microarray.