

ABSTRACT

CANCER DETECTION BASED ON CLASSIFICATION MICROARRAY GENE EXPRESSION DATA USING PRINCIPAL COMPONENT ANALYSIS AND CONJUGATE GRADIENT BACK PROPAGATION

Adiyasa Nurfalalah

Supervisor: Dr. Adiwijaya

Co-Supervisor: Arie Ardiyanti Suryani, S.T., M.T.

World Health Organization states that cancer is the leading cause of death in the world. About 8.2 million people died because cancer and estimated will increase each year due to an unhealthy lifestyle [2]. Deaths due to cancer could be prevented if the cancer was detected early. In recent decades microarray has taken an important role in cancer research. Microarray is a technology that is capable of storing thousands of gene expressions taken from several specific tissues of human at once. By analyzing microarray data, the cancer can be identified from affected genes. This study develops an alternative framework for cancer detection based on microarray data classification using Principal Component Analysis (PCA) and Modified Back Propagation (MBP). MBP is a modification of Standard Back Propagation (BP) by implementing conjugate gradient in training of BP. MBP is better than back propagation because it has fast training. The experiment results by implementing 5-fold cross validation using three public microarray dataset, ovarian cancer data, colon cancer data proved that PCA+MBP requires only around 1-5 second for training. It result was the best result compared with MBP, Back Propagation (BP), and PCA+BP. In terms of accuracy of testing, generally PCA+MBP produce around 76%-97% for all data. It result were not able to outperform PCA+BP that produce accuracy around 83%-98% but MBP was able to achieve the best accuracy with around 83.03%-100% for all data.

Keywords: Cancer, Microarray, Principal Component Analysis, Conjugate Gradient Back Propagation.