

ABSTRACT

Cancer is one of the most common deadly diseases in the world. The conventional diagnostic techniques are not always effective as they rely on the physical and morphological appearance of the tumor. The ability of the physicians to effectively treat and cure cancer is directly dependent on their capability to detect cancers at their earliest stages. Early stage prediction and diagnosis is difficult with those conventional techniques such as physical appearance of tumor. However, these techniques are costly, time consuming, requires large laboratory setup and highly skilled persons. There is need for faster, easier, more accurate and effective method, using modern technology to address the challenges.

In this dissertation, Hybridized model of Genetic Algorithm and Neural Network was developed and simulated in Weka environment using microarray cancer dataset. Microarray studies are characterized by a low sample number and a large feature space with many features irrelevant to the problem being studied. This makes feature selection a necessary pre-processing step for many analyses, particularly classification. Various stages involved in Genetic Algorithm and Neural Network were studied and simulated. An hybrid model that combines the optimization power of Genetic algorithm for reduction of high dimensional microarray data and Neural network for classification between malignant pleural mesothelioma (MPM) and adenocarcinoma (ADCA) of the lung was proposed. The solution found by the combined Genetic Algorithm and Neural Network performed effectively well. The genetic algorithm reduced 12,533 attributes in the microarray dataset to 748 attributes. The reduced microarray dataset was used to train the multilayer perceptron neural network classifier. The trained classifier achieved 97.5% accuracy when evaluated with the testing microarray dataset.

The results presented in this dissertation revealed that the proposed hybrid Genetic Algorithm and Neural Network performs better with over 97% accuracy when used to classify microarray dataset of lung cancer.

Keywords: Cancer, Diagnosis, Genetic Algorithm, Neural Network, Hybridization