ABSTRACT

Data mining is one of the most commonly used techniques in processing the data to get an important information contained in the data. But when confronted by a very large-sized data or high-dimensional, it resulted in the magnitude of the computing resources required to process these data with a fast time. In addition the level ofoptimization and efficiency becomes a problem is also due to the large number of existing attributes.

One of the solutions to do the processing efficiency of high-dimensional data is by applying Evolutionary Algorithm (EAs) into the method of data mining.

Therefore, in this final assignment authors will build a system prediction by applyingone of the algorima EAS i.e. *Genetic Algorithm* to be used in solving the high dimensional data classification. The data used are the data sets obtained from the website of disease Kent Ridge Bio-medical Data Set Repository which each have thousands of attribute data.

Of the system will show the results obtained as someone is potentially exposed to the disease being tested or not. To know the performance of the system that has been built is done the calculation accuracy. So the end result of this experiment is obtained that the use of *Genetic Algorithm* in solving a problem of high-dimensional data classification can be said to be successful. That is because the system has been built producing high enough accuracy values i.e. 89,47% to the data set of the disease of colon tumors and 81.82% to the data set of the disease of leukemia.

Keywords : Classification, Assocaitive Classification, Evolutionary Algorithm, Genetic Algorithm, High Dimensional Data.