

ABSTRACT

Data mining is a method used to process the data with the purpose of obtaining information contained in data. In data mining has many challenges. Examples of challenges in data mining the data contained high dimension. Which means the data has dozens or even thousands of attributes. From the very many attributes that this raises very complex problems for the processing of data.

One solution to solve the problem of high-dimensional data is the incorporation of Data Mining by Evolutionary Algorithm or is often called the Evolutionary Data Mining (EvoDM). The data mining method used is the K-means while the evolutionary algorithm is a momentum-type particle swarm optimization. In order to get the momentum of type hybrid algorithm particle swarm optimization with k-means (MPSO-KM). Hybrid algorithm is used to determine the weight attributes optimal and the cluster center of each attribute.

In this final project built a system to predict the disease with a high-dimensional data in the form of data. The data used was taken from Kent Ridge Biomedical Data Set Repository form of leukemia disease data and Colon Tumors. The accuracy of the prediction data is best obtained with leukemia amounted to 81.36%, while the data for 78.95% of colon tumors.

Keywords: data mining, k-means, momentum-type particle swarm optimization, evolutionary data mining.