

Abstract

Data is a source of useful information for human survival. To make the data useful, we need a method that can dig up important information from existing data. One method of information retrieval of a set of data known as Data Mining. Mine engineering information on Data Mining also varied, one of them *Clustering*. *Clustering* is a method of grouping data that have similar attributes into one group with certain rules. In this research, *Clustering* algorithm used is Density Based Spatial *Clustering* Applications with Noise (DBSCAN). DBSCAN a *Cluster* algorithm that is density-based, ie classifying data based on the density into one group, and data are rare in other groups. To classify high-dimensional data, the device needs to minimize the costs of computing. GPU (Graphics Processing Unit) allows to process high-dimensional data in a short time. If the GPU is combined with DBSCAN grouping data can generate performance algorithms work well with high accuracy and minimum computational cost. One of method applying the GPU on DBSCAN by calculating the distance between the data in parallel on the GPU. The result of this calculation is able to save the cost of computing with time 1.035921875 sec for data with dimensions of 15154 and 0.063893878 seconds for the data to the dimensions of 12600. In addition to the performance evaluation, GPU produces a pretty good value compared with serial algorithm.

Keywords : *Data Mining, Clustering, Density Based Spatial Clustering Application with Noise (DBSCAN) , density based, GPU (Graphics Processing Unit) .*