

ABSTRACT

Data mining is a combination technology for analyze a useful information from dataset using some technique such as classification, clustering, and etc. Clustering is one of the most used data mining technique these day. Its application used as for image segmentation, bioinformatics, and pattern recognition. . K-Means and K-Medoids is one of clustering algorithms that mostly used because it's easy implementation, efficient, and also present good results.

Besides mining important information, the needs of time spent when mining data is also a concern in today era considering the real world applications produce huge volume of data. This research analyzed the result from K-Means and K-Medoids algorithm and time performance using another knowledge area known as High Performance Computing (HPC) Cluster to parallelize K-Means and K-Medoids algorithms and using Message Passing Interface (MPI) library. The results shown than K-Means algorithm gives smaller SSE than K-Medoids. And also parallel algorithm that used MPI gives faster computation time than sequential algorithm.

Keywords: Clustering, K-Means, K-Medoids, HPC Cluster, MPI