

ABSTRAK

Data mining adalah teknologi gabungan untuk menganalisis suatu informasi menggunakan beberapa teknik tertentu seperti klasifikasi, klustering, dll, untuk mengumpulkan informasi berharga dari suatu data set. Klustering merupakan salah satu teknik data mining yang banyak digunakan saat ini seperti untuk *image segmentation*, *bioinformatics*, dan *pattern recognition*. *K-Means* dan *K-Medoids* merupakan algoritma yang banyak digunakan karena kemudahan dalam pengaplikasiannya serta memberikan hasil klustering yang cukup baik.

Selain mendapatkan informasi yang berguna dan akurat, saat ini waktu yang digunakan pada saat memproses data juga ikut menjadi perhatian, mengingat data yang ada saat ini jumlahnya sudah sangat besar. Pada penelitian ini akan di analisis bagaimana hasil klustering menggunakan *K-Means* dan *K-Medoids* serta performansi dari penggunaan *High Performance Computing (HPC) Cluster* untuk memparalelkan algoritma *K-Means* dan *K-Medoids* dengan menggunakan *Message Passing Interface (MPI) library*. Hasil yang didapatkan menunjukkan bahwa algoritma *K-Means* menghasilkan *error* yang lebih kecil dibandingkan algoritma *K-Medoids*. Sedangkan pada proses pengolahan data, waktu komputasi dengan menggunakan MPI menghasilkan waktu komputasi yang lebih cepat dibandingkan waktu komputasi algoritma sekuensial.

Kata kunci: *Clustering, K-Means, K-Medoids, HPC Cluster, MPI*