

## ABSTRAK

Penambangan data adalah proses menemukan pola yang menarik, dan pengetahuan dari data yang berjumlah besar . Ada banyak metode dalam penambangan data dan salah satunya adalah *clustering*,. Metode ini digunakan untuk mengidentifikasi kelompok alami dari sebuah kasus dan mengelompokkan data yang memiliki kemiripan atribut. Nilai daya serap ujian nasional merupakan salah satu data berjumlah besar yang dapat diproses untuk menemukan informasi mengenai kelompok-kelompok Sekolah Menengah Atas di provinsi Daerah Istimewa Yogyakarta (DIY) berdasarkan nilai daya serap ujian nasional .

Dalam tugas akhir ini algoritma *divisive hierarchical clustering* akan diimplementasikan untuk mengelompokkan Sekolah Menengah Atas di provinsi Daerah Istimewa Yogyakarta (DIY) berdasarkan nilai daya serap ujian nasional . Teknik ini merupakan proses pengklasteran yang didasarkan pada persamaan nilai rata-rata antar objek. Proses penambangan dilakukan dengan mengikuti tahapan dalam proses KDD(*Knowledge Discovery in Database*) yaitu pembersihan data integrasi data, pemilihan data, transformasi data, penambangan data, evaluasi pola, dan presentasi pengetahuan .

Berdasarkan evaluasi menggunakan nilai daya serap dari 29 kompetensi dapat disimpulkan bahwa algoritma ini tidak tepat digunakan untuk menganalisis nilai daya serap ujian nasional karena menghasilkan pohon *cluster* yang timpang, namun algoritma ini dapat mengelompokkan nilai rata – rata nilai daya serap meskipun setiap cluster hanya berisi 1 atau 2 anggota saja .

Kata Kunci : Penambangan Data, *Clustering*, *Divisive Hierarchical Clustering*, nilai daya serap.

## *ABSTRACT*

Data mining is the process of finding interesting patterns, and knowledge of large amounts of data. There are many methods in data mining and one of them is clustering. This method is used to identify the natural group of a case and classify data that has similarity attributes. The absorption level of the national examination is one of the large amount of data that can be processed to find information on high school groups in the province of Yogyakarta based on the value of national exam absorption.

In this final project divisive hierarchical clustering algorithm will be implemented to classify High School in Yogyakarta Special Province based on absorption level of national examination. This technique is a clustering process based on the average value equation between objects. The mining process is done by following the steps in the process of KDD (Knowledge Discovery in Database) which is data cleaning data integration, data selection, data transformation, data mining, pattern evaluation, and knowledge presentation.

Based on the evaluation using the absorption level of 29 competencies it can be concluded that this algorithm is not appropriately used to analyze the absorptive value of the national exam because it produces a lame cluster tree, but this algorithm can classify the average value of absorptive power even if each cluster contains 1 or 2 members only.

Keywords: Data Mining, Clustering, Divisive Hierarchical Clustering, absorption level.