

ABSTRAK

Data mining merupakan penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar. Salah satu metode *data mining* yaitu *clustering*. *Clustering* adalah proses mengelompokkan objek ke dalam suatu kelompok (*cluster*) sehingga objek memiliki kemiripan sangat besar dengan objek lain yang berada pada *cluster* yang sama, tetapi memiliki ketidakmiripan yang besar dengan objek yang berada pada *cluster* berbeda. Salah satu algoritma *clustering* adalah *Fuzzy C-Means* (FCM). FCM adalah metode *clustering fuzzy* dimana keberadaan tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. Dengan metode ini, data akan dikelompokkan ke dalam *cluster* yang berbeda sesuai persentase keanggotaannya di masing-masing klaster.

Pada tugas akhir ini akan diimplementasikan algoritma *Fuzzy C-Means* untuk mengelompokkan data Sekolah Menengah Atas di DIY berdasarkan nilai Ujian Nasional dan Nilai Sekolah. Hasil tugas akhir ini yaitu sebuah perangkat lunak yang dapat digunakan sebagai alat bantu untuk pengelompokan Sekolah Menengah Atas menggunakan algoritma *Fuzzy C-Means*. Pengujian terhadap sistem ini terdiri dari pengujian *blackbox*, dan pengujian validitas implementasi algoritma *Fuzzy C-Means*.

Berdasarkan hasil pengujian *blackbox* yang dilakukan, disimpulkan bahwa sistem pengelompokan ini dapat menghasilkan keluaran yang sesuai dengan yang diharapkan pengguna. Berdasarkan hasil pengujian validitas implementasi algoritma *Fuzzy C-Means* yang dilakukan, disimpulkan bahwa Metode *clustering* algoritma *Fuzzy C-Means* terbukti dapat mengelompokkan sekumpulan data bertipe numerik.

Kata kunci: *Data mining*, algoritma *Fuzzy C-Means*, *Clustering*.

ABSTRACT

Data mining is mining process or extracting information by getting pattern or specific rules from large amounts of data . One of the data mining method is clustering. Clustering is the process of grouping objects into a group (cluster) so that the object has a very great similarity with other objects that are on the same cluster, but has a great dissimilarity with objects that are in different clusters. One algorithm of clustering is Fuzzy C-Means (FCM). FCM is a technique of fuzzy clustering in which each data point in a cluster is determined by the degree of membership. With this method, the data will be grouped into different clusters according to percentage of membership in each cluster.

This thesis will implement Fuzzy C-Means algorithm was used to clustering high school data based on the value of National Exam and School Exam. The results of this thesis is a software that can be used as a tool to cluster high school data using Fuzzy C-Means algorithm. The testing of this system consists of blackbox testing and validity testing of Fuzzy C-Means algorithm implementation.

Based on blackbox testing, it can be concluded that the clustering system could produce output as expected and handle any incorrect functions. Based on validity testing of Fuzzy C-Means algorithm implementation, it can be concluded that the clustering method Fuzzy C-Means algorithm, is able to classify the type of numerical data set

Keywords : *Data mining, Fuzzy C-Means algorithm, Clustering.*