ABSTRAK

Analisis *cluster* adalah salah satu teknik analisis statistik yang digunakan untuk meringkas data dengan cara mengelompokkan obyek-obyek berdasarkan kesamaan karakteristik tertentu yang dimiliki masing-masing obyek. Kesamaan karakteristik tersebut dinyatakan dalam ukuran jarak antar obyek. Pembentukan kelompok-kelompok berdasarkan jarak, obyek yang mirip seharusnya berada dalam kelompok yang sama dan mempunyai jarak yang lebih kecil. Sebaliknya obyek yang berbeda berada dalam kelompok yang berbeda dan mempunyai jarak yang lebih besar.

Pada skripsi ini akan dibahas dua metode dalam analisis *cluster*, yaitu metode *hierarki* dan metode *nonhierarki*. Metode *hierarki* dibagi menjadi dua kategori yaitu *agglomerative* dan *divisive*. Metode *nonhierarki* yang akan dibahas pada skripsi ini adalah metode *K-means cluster*.

ABSTRACT

Cluster analysis is one of statistical analysis techniques to summarize data set by grouping objects based on the same specific characteristics of the objects. These characteristics similarity is represented by the distance between objects. The determination of groups is based on the distance of the objects. The identical objects should be placed in the same group and have shorter distance. Conversely, the nonidentical objects should be placed in different groups and have longer distance.

This thesis will discuss two methods of cluster analysis, those are hierarchical method and nonhierarchical method. Hierarchical method is divided into two categories, which are agglomerative and divisive. The only nonhierarchical method which is discussed in this thesis is K-means cluster method.