

ABSTRAK

Nama : Naufal Eldiva VA

NIM : 114010049

Judul : ANALISA KLASSTERING MENGGUNAKAN PARALEL K-MEANS PADA DATA TERJEMAH AL-QUR'AN BAHASA INDONESIA

Salah satu jenis data yang menarik untuk dilakukan data mining adalah teks. Dikenal sebagai text mining, proses data mining pada teks dapat menghasilkan informasi penting yang terkandung dalam teks. Teks yang populer untuk diteliti peneliti muslim adalah Al-Qur'an, hal ini jelas penting karena sebagai umat islam kita diwajibkan untuk membaca dan memahami lebih dalam isi Al-Qur'an. Untuk mempermudah memahami isi Al-Qur'an kita dapat mempelajari terjemahannya. Salah satu proses teks mining yang populer dan penting penggunaannya adalah klasterisasi. Terdapat banyak metode klasterisasi, salah satunya adalah K-means. Pada penelitian ini dilakukan percobaan klasterisasi menggunakan metode simple K-means dan parallel K-means dengan bantuan program python yang berbasis online. Sehingga dapat dilihat hasil dari tiap metode dan perbandingannya. Jumlah kluster pada penelitian ini dipilih secara acak yakni 2, 3, 4, 5, 10, 20, 50, 100, 114. Berdasarkan jumlah kluster tersebut maka waktu yang dibutuhkan simple K-means adalah 0.22 detik sampai 10.5 detik dan pada parallel K-means adalah 0.23 detik sampai 10.3 detik. Parallel K-means lebih efektif saat jumlah kluster ≥ 10 namun membutuhkan waktu yang sama saat jumlah kluster < 10 bahkan lebih lambat saat jumlah kluster 5. Adapun urutan terdekat pada centroid dan isi kluster metode simple K-means dan parallel K-means sama atau hanya berbeda urutan item pada tiap pemilihan kluster kecuali pada saat jumlah kluster 114 yang menghasilkan urutan dan isi item yang jauh berbeda tiap metodenya.

Kata Kunci: Klastering, *K-means*, *Parallel K-means*, Klastering Al-quran Bahasa indonesia.

ABSTRACT

Name : Naufal Eldiva V.A

NIM : 1147010049

Title : CLUSTERING ANALYSIS USING PARALLEL K-MEANS ON INDONESIAN QUR'AN TRANSLATION DATA

One type of data that is interesting for data mining is text. Known as text mining, the process of data mining on text can produce important information contained in the text. A popular text for Muslim researchers to study is the Qur'an, this is clearly important because as Muslims we are required to read and understand more deeply the contents of the Qur'an. To make it easier to understand the contents of the Qur'an we can study its translation . One of the popular and important text mining processes is clustering. There are many clustering methods, one of which is K-means. In this study, clustering experiments were carried out using the simple K-means and parallel K-means methods with the help of an online-based python program. So that it can be seen the results of each method and its comparison. The number of clusters in this study was chosen randomly, namely 2, 3, 4, 5, 10, 20, 50, 100, 114. Based on the number of clusters, the time required for simple K-means is 0.22 seconds to 10.5 seconds and in parallel K- means are 0.23 seconds to 10.3 seconds. Parallel K-means is more effective when the number of clusters is 10 but takes the same time when the number of clusters <10 is even slower when the number of clusters is 5. The closest order to the centroid and the contents of the clusters of the simple K-means and parallel K-means methods are the same or only different order of items in each cluster selection except when the number of clusters is 114 which results in the order and content of items that are much different for each method.

Keywords: Clustering, *K-means*, *Parallel K-means*, Indonesian Al-quran Clustering.