

INTISARI

INISIALISASI PUSAT KLASTER (*CENTROID*) AWAL PADA K-MEANS DENGAN METODE CENTRONIT

Oleh

Maria Silvia

12/334609/PA/14842

K-Means merupakan salah satu algoritma analisis klaster yang sangat terkenal karena kemampuannya untuk mengelompokkan data besar dengan sangat cepat. Akan tetapi performa K-Means sangat bergantung pada inisialisasi pusat klaster (*centroid*) awal. Biasanya inisialisasi pusat klaster awal pada K-Means ditentukan secara acak. Perbedaan pada pusat klaster awal akan memberikan perbedaan pula pada hasil klaster. Apabila pusat klaster awal yang diberikan adalah pusat klaster yang baik (pusat klaster yang dapat merepresentasikan anggota-anggota dalam suatu klaster), maka hasil pengklasteran dapat dipastikan juga baik. Namun hal tersebut tidak selalu terjadi, sehingga dapat dipastikan hasil klaster yang terbentuk berubah-ubah (tidak konsisten).

Centronit merupakan pengembangan metode inisialisasi pusat klaster awal pada K-means. Algoritma pada metode ini berdasarkan pada nilai minimum rata-rata perhitungan jarak antar data. Metode ini juga *robust* terhadap data outlier. Dengan metode centronit, inisialisasi pusat klaster awal dapat ditentukan terlebih dahulu sehingga dapat dihasilkan hasil pembentukan klasternya yang lebih baik dan konsisten.

Kata kunci: K-Means, Analisis klaster, Inisialisasi pusat klaster awal, Centronit

ABSTRACT

INITIALIZATION CENTROID FOR K-MEANS USING CENTRONIT METHOD

by

Maria Silvia

12/334609/PA/14842

K-Means is one of very well known clustering algorithm because its ability to classify large data very quickly. However, clustering performance of the K-Means highly depends on the initial centroids. Usually initial centroids for the K-Means clustering are determined randomly. Difference in the initial centroid will make a difference in the results of clustering. If the initial given centroid is good (centroid that can represent the members in a cluster), then the clustering results can be ensured also good. Yet, it does not always happen, so it can be ensured that its cluster formation changeable (inconsistent).

Centronit is an initialization method development of the initial centroids in K-Means. The algorithm of this method is based on the calculation of the average distance of the nearest data use the minimum distance. This method also robust from outliers of data. By using centronit method, the initial centroids can be determined in advance so it can produce the better and consistent cluster formation.

Key words: K-Means, Clustering, Initial centroids, Centronit