

ABSTRAK

Nama : Siti Mahfuzdoh
NIM : 1167010066
Judul : *Modified Union* Ekstraksi Fitur untuk Algoritma CLARANS pada Data Terjemahan Hadis

Berkembangnya teknologi yang semakin meningkat mengakibatkan penumpukan data dalam jumlah yang besar, penumpukan data ini juga terjadi pada data hadis. Algoritma *clustering* dibutuhkan untuk mengatasi permasalahan ini. Pada penelitian ini digunakan algoritma *clustering* CLARANS yang efektif pada pengelompokan data dalam jumlah yang besar. Untuk meningkatkan kinerja algoritma *clustering* maka dilakukan tahap reduksi dimensi fitur untuk mengurangi fitur-fitur yang kurang berpengaruh dan redundan. Metode penggabungan *modified union* diterapkan untuk menggabungkan metode *document frequency* dan *term variance*, kemudian data hasil reduksi dari metode *modified union* direduksi kembali menggunakan PCA pada *feature extraction*. Pada penelitian ini dilakukan perbandingan hasil *clustering* pada data terjemahan hadis yang berjumlah 2175 hadis dengan menggunakan algoritma CLARANS sebelum dan sesudah dilakukan reduksi untuk mengetahui hasil *clustering* yang terbaik pada kedua percobaan tersebut. Hasil *clustering* yang didapat setelah dilakukan evaluasi menggunakan *Davies-Bouldin Index* (DBI) dan *Silhouette Coefficient* (SH) menunjukkan *clustering* algoritma CLARANS setelah dilakukan reduksi mempunyai hasil yang lebih baik jika dibandingkan dengan hasil algoritma *clustering* CLARANS sebelum dilakukan reduksi dengan nilai DB minimal 1.7738 dan SC maksimal 0.0175. Kemudian algoritma CLARANS juga efektif untuk mengelompokkan kumpulan data besar.

Kata Kunci: *Modified Union*, *Feature Extraction*, PCA, *Davies-Bouldin Index*, *Silhouette Coefficient*

ABSTRACT

Name : Siti Mahfuzdoh

NIM : 1167010066

Title : *Modified Union Feature Extraction for the CLARANS Algorithm
on Hadith Translation Data*

The development of technology which is increasingly increasing resulted in the accumulation of large amounts of data, this accumulation of data also occurs in the data of hadith. Clustering algorithm is needed to solve this problem. In this study, CLARANS clustering algorithm is used which is effective in grouping large amounts of data. To improve the performance of the clustering algorithm, a feature dimension reduction step is performed to reduce features that are less influential and low. The modified union merging method is applied to combine the document frequency and term variance methods, then the data resulting from the modified union method is reduced again using PCA in the extraction feature. In this study a comparison of the results of clustering on the hadith translation data totaling 2175 hadith using the CLARANS algorithm before and after the reduction is carried out to determine the best clustering results in the two experiments. Clustering results obtained after evaluation using the Davies-Bouldin Index (DBI) and Silhouette Coefficient (SH) show that CLARANS algorithm clustering after the reduction has better results compared to the results of the CLARANS clustering algorithm before the reduction with a minimum DB value of 1.7738 and SC a maximum of 0.0175. Then the CLARANS algorithm is also effective for grouping large data sets.

Keywords : *Modified Union, Feature Extraction, PCA, Davies-Bouldin Index, Silhouette Coefficient*