

INTISARI

ANALISIS *WEB USAGE MINING* UNTUK PEMBENTUKAN PROFIL *USER* MENGGUNAKAN ALGORITMA *CLUSTERING K-MEANS* (STUDI KASUS: *WEBSITE ETD.UGM.AC.ID*)

Oleh

Wildan Aulia Hikmawan

13/34/7518/PA/15280

Perkembangan teknologi internet saat ini sangat pesat dapat dilihat dari kemudahan mendapatkan akses internet. Seiring dengan hal tersebut, terjadi peningkatan aktivitas pengguna dalam mengunjungi berbagai *website*.

Etd.ugm.ac.id merupakan salah satu *website* di Universitas Gadjah Mada. *website* tersebut dipilih sebagai bahan penelitian karena *website etd.ugm.ac.id* cukup sering diakses oleh mahasiswa untuk mencari referensi penelitian atau tugas akhir. Pola akses pengunjung *website* tersebut dapat digunakan untuk mengetahui kebiasaan seorang pengunjung dalam melakukan penelusuran dalam sebuah *website*. Hal tersebut dapat digunakan guna perbaikan kualitas *website* dan meningkatkan kepuasan pengunjung dalam mengakses *website* untuk menemukan informasi di dalamnya sesuai dengan kebutuhannya.

Pada penelitian ini dilakukan pengelompokan pengunjung berdasarkan log akses dan konten yang diakses, dengan menggunakan algoritma K-Means. Pengujian sistem dilakukan dengan menghitung jarak antar cluster yang terbentuk dan jarak antar anggota pada suatu cluster dengan metode silhouette coefficients. Sehingga didapatkan nilai silhouette coefficient 0,05 untuk cluster dokumen dengan 500 fitur kata dan $K = 30$, sedangkan untuk cluster pola akses buku oleh pengguna mendapatkan nilai silhouette coefficient 0,15 dengan jumlah cluster 14.

Kata-kata kunci : *website*, profil navigasi, clustering, pemodelan topik, pattern mining

ABSTRACT

WEB USAGE MINING ANALYSIS FOR CREATING USER PROFILE USING K-MEANS CLUSTERING ALGORITHM (CASE STUDY: *ETD.UGM.AC.ID WEBSITE*)

by

Wildan Aulia Hikmawan

13/34/7518/PA/15280

The development of internet technology today is very rapid can be seen from the ease of getting internet access. Along with this, there is an increase in user activity in visiting various websites.

Etd.ugm.ac.id is one of the websites at Gadjah Mada University. It was chosen as research material because website etd.ugm.ac.id quite often accessed by students to search for research reference or final project. Visitor pattern of the website can be used to know the habits of a visitor in doing a search in a website. It can be used to improve the quality of the website and increase the satisfaction of visitors in accessing the website to find information in it according to their needs.

In this study, visitors are grouped based on access logs and accessed content, using K-Means algorithm. System testing is done by calculating the distance between the clusters formed and the distance between members on a cluster with the method of silhouette coefficients. So the value of silhouette coefficient 0,05 for document cluster with 500 word feature and $K = 30$, whereas for cluster access pattern book by user get value of silhouette coefficient 0,15 with number of cluster 14.

Keywords: *website*, navigation profile, clustering, topic modeling, mining patterns