

BAB I PENDAHULUAN

1.1 Latar Belakang

Selama era teknologi yang semakin berkembang, aktivitas yang terhubung ke dunia internet sudah menjadi kebutuhan setiap orang. Pengguna mengumpulkan lebih banyak data saat berinteraksi dengan situs web, media sosial, aplikasi, dan layanan online lainnya. Informasi tentang perilaku, preferensi, dan kebiasaan pengguna dalam berbagai situasi dimasukkan ke dalam data ini. Cyber profiling penelitian biasanya menggunakan eksplorasi data untuk mengetahui aktifitas pengguna saat mengakses konten web. Algoritma K-MEANS, yang dapat membagi data berdasarkan jumlah pengunjung terbanyak, adalah salah satu cara yang mungkin bermanfaat proses profiling. Namun, dalam studi ini, kategori dibuat secara manual berdasarkan konten website dan waktu akses. Tujuan grup ini adalah untuk melihat semua web yang pengguna sering akses.

Menurut Diana dan Apridiansyah (2019), File yang dikenal sebagai access log berisi semua tindakan akses ke server. Analisis akses log diterapkan untuk mengumpulkan data yang disimpan di server. Data yang digunakan berasal dari server web resmi kampus polinela. Website Polinela.ac.id berfungsi untuk menyediakan informasi yang bermanfaat, komunikatif, dan edukatif bagi masyarakat umum, mahasiswa/i, dan pendidik. File yang biasanya disimpan di server log mencatat setiap kunjungan ke situs web ini. File ini menyimpan informasi seperti alamat IP, waktu, halaman yang diakses, dan informasi lainnya. Jika server log situs web dibiarkan begitu saja, mereka hanya akan bertambah banyak sehingga memenuhi penyimpanan server. Diperlukan tindakan, seperti menggunakan web mining. Algoritma cluster, dapat digunakan untuk menentukan kelompok objek yang bermanfaat. Digunakan pada saat tujuan analisis data. Profiling juga harus digunakan untuk mengidentifikasi ciri pengguna dan akses yang dicoba. Oleh karena itu,

profiling akan sangat membantu mengetahui sikap pengguna saat mengunjungi situs web institusi pendidikan di Indonesia. waktu ketika kunjungan internet dapat muncul dan terjadi kegiatan yang mencurigakan, terutama karena tidak sering ada gejala yang menunjukkan bahwa peserta didik terlibat dalam cybercrime. Untuk alasan ini, penting untuk mempelajari lebih lanjut tentang bagaimana kegiatan kunjungan online yang sering digunakan di dunia pembelajaran sesuai dengan lingkup kegiatan di dunia pembelajaran. Salah satu cara untuk mengetahui sikap pengunjung website adalah dengan menerapkan proses profiling yang bergantung pada log akses server website. Karena itu, hasil pengujian ini untuk dapat mengidentifikasi aktivitas apa yang dilakukan pengguna saat mengakses website dan profil pengguna. maka dari itu penulis bermaksud merancang sistem analisis yang akan dijadikan sebuah Laporan Akhir dengan judul” ANALISIS ACCESS LOG WEB MENGGUNAKAN K-MEANS CLUSTERING UNTUK CYBER PROFILING (Studi Kasus Politeknik Negeri Lampung)”

1.2 Identifikasi Masalah

Penelitian web saat ini sudah umum. Seringkali, orang hanya berinteraksi melalui internet, dan identitas mereka bergantung pada perilaku yang mereka tunjukkan di internet. Untuk menerapkan metode pembuatan profil, Anda perlu mendapatkan akses ke perilaku pengguna. Pembuatan profil menggunakan log akses web server diharapkan mengetahui profil pengguna saat melakukan penelusuran website secara online.

1.3 Rumusan Masalah

Dengan demikian, rumusan masalah penelitian ini adalah sebagai berikut:

1. Apakah algoritma K-MEANS Clustering dapat digunakan dengan melakukan proses cyber profiling dengan mengkategorikan data access log web server berdasarkan website yang menerima paling banyak kunjungan?
2. Bagaimana K-MEANS Clustering dapat mengidentifikasi profil pengguna dalam penelusuran konten web?

1.4 Ruang Lingkup Penelitian

Ruang lingkup dari pelaksanaan penelitian atau sasaran yang akan dicapai dengan melakukan pengelompokan data pengunjung website polinela.ac.id melalui objek data access log web server, dikampus politeknik negeri lampung. Penelitian diharapkan dapat mengumpulkan data tentang kunjungan web seperti aktifitas pengguna kunjungan halaman web dan data pengguna yang mengakses waktu kunjungan web.

1.5 Batasan Masalah

1. Fokus penelitian ini hanyalah data log akses web server pada kampus politeknik lampung.
2. Fokus Studi ini terbatas pada analisis profil pengunjung website berdasarkan jumlah pengunjung terbanyak dan waktu akses.
3. Studi ini menggunakan clustering K-MEANS dengan jupyter notebook.

1.6 Tujuan Penelitian

Berikut ini adalah tujuan penelitian yang diharapkan dan ingin dicapai:

1. Mengkategorikan konten web server berdasarkan jumlah kunjungan terbanyak.
2. Uji hasil K-MEANS untuk menentukan ciri-ciri pengguna untuk melakukan aktivitas penelusuran web.

1.7 Manfaat Penelitian

Studi ini bertujuan untuk menemukan aktivitas yang sering dilakukan pengunjung web dan profil pengguna layanan web yang ada di politeknik di salah satu perguruan tinggi Lampung.

1.8 Sistematika Penulisan

Adapun dalam penulisan skripsi ini terdiri dari lima bab dengan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini terdiri latar belakang masalah, rumusan masalah, ruang lingkup penelitian, tujuan penelitian, batasan penelitian, manfaat penelitian, metode penulisan sistematis.

BAB II TINJAUAN PUSTAKA

Teori-teori yang mendukung penelitian penulis disajikan dalam bab ini.

BAB III METODOLOGI PENELITIAN

Pada Bab ini memberikan penjelasan tentang metode penelitian penyelesaian masalah dan rumusan masalah. Penjelasan ini mencakup metode pengumpulan data, metode pengembangan sistem, dan alat dan bahan pendukung.

BAB IV HASIL DAN PEMBAHASAN

Hasil dari penelitian, yang mencakup iterasi data secara berkala menggunakan perhitungan Teknik Euclidean dalam Algoritma K-MEANS, akan dibahas pada bab ini.

BAB V PENUTUP

Merangkum hasil yang diperoleh dari pembahasan bab sebelumnya dan menyarankan cara-cara yang dapat digunakan untuk membangun dan mengembangkan penelitian ini di masa mendatang.

DAFTAR PUSTAKA

LAMPIRAN