

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pada era globalisasi modern ini, internet menjadi sebuah komponen yang tidak dapat dipisahkan oleh manusia. Hampir setiap harinya manusia mengakses internet untuk memenuhi kebutuhan pribadinya sebagai contoh penggunaan transportasi dengan fasilitas *Go-Jek*. Dengan kata lain, hampir semua *stakeholder* (baik yang berbasis keuntungan / tidak) mempromosikan dirinya melalui *website*. Promosi ini diharapkan dapat meningkatkan keuntungan bagi *stakeholder* tersebut (baik dalam bentuk materiil / bukan).

Salah satu ukuran kesuksesan dari sebuah *website* adalah tingginya jumlah pengunjung. Namun, *trend* jumlah pengunjung tidak dapat diprediksi dengan mudah karena *stakeholder* sulit untuk mengetahui kapan seorang pengguna akan pergi meninggalkan sebuah *website*. Kepergian dari pengguna dapat mempengaruhi keuntungan dari *stakeholder*. Sebagai contoh, *Friendster.com* yang awalnya merupakan *social media* yang banyak digemari oleh pengguna, kini sudah tutup dikarenakan banyak pengguna yang telah meninggalkannya. *Friendster* sendiri sudah mencoba untuk mencoba melakukan inovasi baru untuk membuat penggunaanya kembali, namun semuanya telah terlambat dan pada akhirnya *Friendster* memutuskan untuk menutup *website*-nya. [1]

Oleh karena itu, perlu adanya sebuah pendeteksi yang dapat membantu *stakeholder* dalam mengetahui aktivitas pengguna. Dengan adanya sistem tersebut diharapkan dapat membantu *stakeholder* untuk menangani kepergian dari pengguna, sebagai contoh apabila *stakeholder* telah mengetahui pengguna akan meninggalkannya, *Stakeholder* dapat memberikan sebuah penanganan ataupun sebuah fitur terakhir untuk pengguna.

Penelitian ini akan terfokus pada sebuah pembuatan sistem prediksi untuk mengetahui pengguna akan meninggalkan sebuah *website* atau tidak. Sistem prediksi ini didapatkan dengan membuat fitur-fitur dari hasil observasi terhadap *web dan software behavior log*. *Log* tersebut memungkinkan untuk didapatkan

dengan menanamkan sebuah program kecil untuk mengetahui setiap aktivitas yang dilakukan oleh pengguna namun, penanaman program tersebut dilakukan atas persetujuan dari pengguna. Namun, pada penelitian ini, hal tersebut tidak akan dibahas dan hanya terfokus terhadap bagaimana *web dan software behavior log* apakah dapat memberikan dampak dan pengaruh dari sistem prediksi tersebut. Sistem prediksi akan didasarkan pada algoritma *multinomial naïve-bayes*.

1.2 Rumusan Masalah

Berdasarkan masalah yang didapat dari kesulitan *Stakeholder* dalam mengetahui kepergian pengguna terhadap suatu *website*, maka dapat dirumuskan rumusan masalah menjadi sebagai berikut:

1. Bagaimana cara mengetahui bahwa seseorang tidak akan mengakses kembali sebuah *website*?
2. Seberapa akurat *web dan software behavior log* dapat memprediksi konsistensi akses *website*?

1.3 Tujuan Pembahasan

Berdasarkan rumusan masalah yang ada pada subbbab 1.2, terdapat beberapa tujuan pembahasan sebagai berikut:

1. Membangun sebuah model prediksi yang didasarkan dari *web dan software behavior log*.
2. Mengevaluasi atribut-atribut yang didapat dari *software dan web behavior log* dengan *k-fold cross validation*.

1.4 Ruang Lingkup

Ruang lingkup tugas akhir ini terfokus untuk menguji dampak *web dan software log behavior* dalam memprediksi konsistensi akses sebuah *website*. Penelitian ini tidak akan membahas bagaimana cara mendapatkan *web dan software log behavior*. Dalam penelitian ini, *data set* diperoleh dari manusia.

1.5 Sumber Data

Sumber data tugas akhir ini bersumber pada:

1. Jurnal, buku, dan halaman web yang berhubungan dengan sistem prediksi, *classification*, *naïve-bayes*, dan *k-fold cross validation*.
2. Pengamatan data *web dan software log behavior*.

1.6 Sistematika Penyajian

Sistematika pembahasan yang akan digunakan dalam laporan ini adalah sebagai berikut:

BAB I. PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, tujuan pembahasan, ruang lingkup, sumber data, dan sistematika penyajian dari tugas akhir ini.

BAB II. KAJIAN TEORI

Bab ini akan menjelaskan teori yang berhubungan dengan *classification*, *k-fold cross validation* dan *naïve-bayes*,

BAB III. ANALISIS DAN RANCANGAN

Bab ini akan berisi *pseudocode*, *use-case* serta *activity diagram* yang akan menggambarkan ide dari pembuatan sistem prediksi.

BAB IV. IMPLEMENTASI

Bab ini akan berisi *data pre-processing* dimana didalamnya terdapat *sub-bab* untuk setiap pemrosesan data. Selain itu juga berisi implementasi *pseudocode* menjadi kode dan prediksi konsistensi akses *web*.

BAB V. PENGUJIAN

Bab ini akan berisi pengujian dimana dalam penelitian ini menggunakan *k-fold cross validation* dimana *k* nya bernilai 10.

BAB VI. KESIMPULAN

Bab ini akan berisi hasil statistik dan akurasi dari setiap fitur.

