

BAB 1 PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika pembahasan yang digunakan dalam Implementasi Naive Bayes Pada Proxy Server Untuk Klasifikasi Pengguna Internet– *Naive Bayes Algorithm*.

1.1 Latar belakang

Penggunaan akses jaringan internet di seluruh dunia mempunyai perkembangan yang sangat pesat. Informasi yang dicari dapat ditemukan dengan sangat mudah. keberadaan internet pun bisa berdampak positif dan sekaligus bisa berdampak negatif bagi penggunaannya (Azwar Rhosyied, 2005).

Menurut pengumpulan data yang diperoleh dari *Asosiasi Penyelenggara Jasa Internet Indonesia* (APJII) sejak tahun 2013 sampai 2014, penetrasi angka pengguna di Indonesia terus mengalami peningkatan yang tinggi sebesar 71.19 juta penggunaan internet di Indonesia pada awal 2013 hingga dengan peningkatan sebesar 28% pada tahun 2014. Pengguna internet di Indonesia paling banyak diakses di Indonesia bagian barat, yakni di pulau Jawa (terutama di kota-kota besar seperti Jakarta dan Surabaya), Bali dan Sumatera (Indonesia Investment, 2013). Hal ini berpengaruh pada perilaku setiap orang yang menggunakan jaringan internet di Indonesia.

Pemerintah saat ini dalam melakukan upaya pembatasan akses pada setiap jaringan internet yang berada di seluruh Indonesia dengan membangun program internet positif pada Juli 2014, dalam rangka mengantisipasi *website* ilegal yang bertentangan dengan *Undang-Undang Informasi dan Transaksi Elektronik* (UU ITE) pasal 27 tentang konten ilegal, yang menangani situs bermuatan negatif yang di sahkan oleh *Menteri Komunikasi Dan Informasi* (MenKomInfo) dalam memblokir sejumlah *Uniform Resource Locator* (URL) yang dianggap tidak pantas dalam menampilkan informasi yang merugikan penggunaannya.

Dengan adanya pembatasan akses yang dilakukan oleh pemerintah, diperlukan suatu sistem yang dapat menganalisa kriteria /jumlah perkembangan kriteria penggunaannya dengan cara mengumpulkan semua aktivitas informasi yang diakses oleh penggunaannya, sehingga dapat mengetahui *Website* mana yang harus di blokir agar secara mudah dapat memblokir sebuah *website* yang mengandung konten Negatif.

Pendekatan dalam pengambilan *Capture Data* pada aktivitas pengguna dilakukan dengan menyisipkan perangkat *Router Mikrotik* pada suatu jaringan internet dengan menggunakan squid web proxy yang berfungsi sebagai aplikasi untuk merekam seluruh kegiatan browsing pengguna, sehingga data aktivitas pengguna internet dapat dikumpulkan.

Selanjutnya penentuan kategori pada suatu *website* diambil dengan metode *Naive Bayes*. Ide dasar *Naive Bayes* adalah *machine learning* yang menggunakan perhitungan probabilitas data yang ada, dalam arti dapat melakukan pengelompokan pada data data yang di dapat. Pada *Naive Bayes Classifier* (NBC) dapat membentuk tabel probabilitas sebagai dasar proses klasifikasi yang dapat memprediksi dan memberikan rekomendasi untuk proses yang di olah dengan nilai optimal. Hasil pengujian akan berpengaruh dengan data klasifikasi yang dapat untuk menentukan besar keberhasilan dari klasifikasinya. Hasil dari penentuan keputusan dapat tingkat accuracy berbeda beda mulai dari 50%, 70% dan 85% dengan nilai maksimal 92.9%. Untuk menghindari Tingkat accuracy yang rendah, maka diperlukan

penyesuaian dari data latih yang dikumpulkan berdasarkan *Kamus Besar Bahasa Indonesia*(KBBI).

Berdasarkan dari latar belakang yang telah dipaparkan, dapat di simpulkan bahwa metode *naive bayes* merupakan salah satu metode yang dapat melakukan pengelompokan jenis karakter pengguna internet dengan mengambil data melalui squid server, maka dalam proposal skripsi ini diusulkan dengan **“Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internet– Naive Bayes Algorithm”**. Diharapkan dengan menggunakan metode *naive bayes* dapat mengetahui tingkat jenis kriteria analisis pengguna internet tersebut.

1.2 Rumusan masalah

Berdasarkan pembahasan latar belakang yang telah diuraikan , maka dapat dirumuskan beberapa rumusan masalah sebagai berikut:

1. Bagaimana cara kerja Squid server dalam mengumpulkan data *Uniform Resource Locator* (URL)?
2. Bagaimana menerapkan metode *Naive Bayes* dengan menghubungkan data yang di kumpulkan oleh squid server untuk mengetahui kriteria pengguna internet?
3. Bagaimana mengetahui cara kerja pemblokiran *Uniform Resource Locator* (URL) yang dilakukan oleh Sistem ?
4. Bagaimana mengetahui kekurangan dan kelebihan pada System yang berjalan?

1.3 Tujuan

Tujuan dari penelitian yang dibahas adalah sebagai berikut :

1. menerapkan metode *Naive Bayes* dengan menghubungkan data yang di kumpulkan untuk mengetahui Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internet.
2. Mengetahui tingkat perkembangan Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internet– *Naive Bayes Algorithm*.

1.4 Manfaat

Dalam penulisan ini diharapkan akan memberikan manfaat yang besar bagi penulis, pembaca, serta masyarakat umum. Dengan arti antara lain:

1. Bagi penulis
 - a. Sebagai hubungan media pembelajaran untuk mengimplemetasikan metode *naive bayes* pada analisis kriteria penggunaan internet di indonesia.
 - b. Mengetahui tingkat perkembangan analisis kriteria pengguna internet menggunakan alogritma naive bayes pada proxy server.
2. Bagi pembaca
 - a. Mendapat pengetahuan akan cara kerja metode *Naive Bayes*.
 - b. Mendapatkan infomasi mengenai kriteria setiap pengguna.
3. Bagi pemerintah
 - a. Mendapatkan semua hasil dari kriteria pengguna yang menggunakan jaringan internet di indonesia.
 - b. Mengurangi besar resiko kesalahan dalam pengambilan keputusan akbiat kurangnya mengenal sifat karakteritik pengguna yang ada di indoensia.

1.5 Batasan masalah

Untuk dapat memfokuskan pada tujuan penelitian, maka penulis akan membatasi ruang lingkup dari penelitian ini. Adapun yang akan menjadi keputusan ruang lingkup pada penulisan ini adalah :

1. Kriteria data pengelompokan pada pengguna internet dibatasi sampai dengan 10-150 data saja.
2. Pengujian dalam mengetahui tingkat perkembangan analisis kriteria penggunaan internet masih dilakukan dalam lingkungan sekitar kampus.

1.6 Sistematika pembahasan

Sistematika laporan dalam penelitian ini dibagi menjadi tujuh bab masing masing diuraikan sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan mengenai latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian, dan sistematika laporan yang di gunakan dalam penyusunan proposal skripsi Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internet- *Naive Bayes Algorithm*.

BAB II LANDASAN KEPUSTAKAAN

Pada bab ini menhelaskan mengenai dasar dasar teori yang digunakan sebagai dasar untuk menyelesaikan permasalahan yang ada. Dasar teori yang diperlukan untuk mendukung penelitian ini adalah karakteristik Pengguna, *Squid Server, Regular Expression, Pre Proccesing, Naive Bayes Algorithm*.

BAB III METODOLOGI

Pada bab ini menjelaskan tentang metode penelitian yang akan digunakan dalam Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internet- *Naive Bayes Algorithm*. Beberapa hal yang dibahas dalam bab ini adalah study litelatur, analisis kebutuhan, pengumpulan data, identifikasi masalah, proses klasifikasi, perancangan sistem, implementasi, serta pengujian dan analisis yang digunakan dalam penelitian.

BAB IV PERANCANGAN SISTEM

Pada bab ini menjelaskan tentang rencana dari Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internetyang akan dibangun.

BAB V IMPLEMTASI

Pada bab ini menhelaskan tentang implementasi dari Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internet- *Naive Bayes Algorithm*.

BAB VI HASIL DAN PEMBAHASAN

Pada bab ini dijelaskan mengenai Implementasi Naive Byes Pada Proxy Server Untuk Klasifikasi Pengguna Internet- *Naive Bayes Algorithm*. Pada bab ini juga dibahas mengenai pengujian dalam mengetahui tingkat perkembangan analisis kriteria penggunaan internet.

BAB VII PENUTUP

Pada bab ini dijelaskan mengenai ringkasan dan kesimpulan dari pengujian terhadap Implementasi Naive Bayes Pada Proxy Server Untuk Klasifikasi Pengguna Internet- *Naive Bayes Algorithm*.