

BAB I

PENDAHULUAN

1.1 Latar Belakang

Saat ini teknologi informasi dan komunikasi semakin berkembang, SMS atau *Short Message Service* merupakan hal yang sudah dikenal oleh masyarakat di dunia. Dengan SMS, komunikasi antara ponsel dapat dilakukan di mana saja, kapan saja dan oleh siapa saja. Karena sifatnya yang umum, maka SMS seringkali digunakan sebagai perantara untuk menyebarkan *spam* sesuai dengan kepentingan masing-masing individu.

Secara definisi SMS *spam* adalah pesan teks yang tidak diinginkan atau tidak diminta yang dikirim secara berturut-turut. Namun pada saat ini SMS *spam* secara umum dapat dikatakan sebagai pesan teks yang tidak diinginkan karena biasanya hanya dikirim sebanyak satu kali ke setiap nomor telepon. SMS *spam* yang dikirim dapat berupa iklan atau promosi ataupun sebuah pesan penipuan yang mengatakan anda mendapatkan hadiah dengan mengatasmakan perusahaan atau lembaga tertentu. Beberapa orang mungkin akan tergiur dengan hal tersebut sehingga mencoba menghubungi *spammer* dan akhirnya tertipu. Sebagian orang yang tertipu mungkin merasa rugi karena telah mengeluarkan uang yang ditransfer ke *spammer* sebagai syarat pengiriman hadiah.

Di Indonesia banyak terjadi kasus penipuan lewat SMS, salah satunya dikutip dari dektiknews pada 2 Maret 2021, dua pelaku penipuan via sms yang meraup 200 juta sebulan dibekuk polisi. Kedua pelaku tersebut menipu dengan modus undian dengan mengirimkan pesan secara acak bahwa penerima pesan memenangkan undian, kemudian akan dijelaskan oleh pelaku bagaimana cara mendapatkan hadiahnya yakni dengan mentrasfer sejumlah uang sebagai persyaratan. Selain itu pada tahun 2019 sebuah aplikasi mobile yang bernama Truecaller merilis daftar negara dengan sms spam terbanyak dan Indonesia termasuk di dalamnya pada urutan ke 10. Untuk mengatasi masalah tersebut dapat digunakan salah satu cara yaitu dengan menggunakan teknik *Machine Learning* yang dapat mengklasifikasi SMS apakah itu SMS normal, SMS promo ataupun SMS penipuan. Selain menggunakan *Machine Learning*, juga dibutuhkan teknik Text Mining untuk memproses data sebelum dilakukannya proses *machine learning* agar data menjadi efisien.

Untuk melakukan klasifikasi dalam penelitian ini, algoritma yang digunakan adalah *Naïve Bayes Classifier*. Naive Bayes merupakan suatu metode *Machine Learning* yang populer dan sering digunakan untuk melakukan klasifikasi, selain itu algoritma ini mudah diimplementasikan, memiliki kinerja yang baik, dan prosesnya cepat [1]. Kemudian dari segi akurasi tidak jauh berbeda dari algoritma lain seperti SVM (*Support Vector Machine*) dan *Logistic Regression* [2], oleh karena itulah algoritma Naïve Bayes digunakan pada penelitian ini.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut :

1. Berapa akurasi dari Naïve Bayes Classifier dalam menentukan jenis SMS spam ?
2. Bagaimana cara mengklasifikasikan SMS *spam* berbahasa Indonesia dengan menggunakan algoritma Naïve Bayes Classifier ?

1.3 Batasan Masalah

Batasan-batasan masalah dalam penulisan skripsi ini antara lain :

1. Objek yang diklasifikasikan adalah SMS berbahasa Indonesia.
2. Jenis SMS yang diklasifikasikan dibagi menjadi tiga kategori atau kelas yaitu normal, penipuan, dan promo.
3. SMS yang akan diklasifikasi harus melalui tahap *preprocessing* terlebih dahulu.
4. Membuat model untuk melakukan pengujian yaitu mengidentifikasi jenis SMS yang diinput.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui hasil akurasi dari Naïve Bayes Classifier dalam mengklasifikasi sms spam berbahasa Indonesia dan dapat menentukan jenis SMS apakah termasuk jenis normal, penipuan, atau promo. Selain itu menguji apakah Bernoulli Naïve Bayes dan Gaussian Naïve Bayes dapat mengklasifikasikan data teks dengan banyak kelas dan yang sifatnya tidak kontinu.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini yaitu mengetahui tipe dari Naïve Bayes Classifier yang dapat melakukan klasifikasi terhadap SMS spam beserta kinerja dari setiap tipenya.

1.6 Sistematika Penulisan

Adapun sistematika penulisan skripsi ini adalah sebagai berikut :

BAB I PENDAHULUAN

Membuat kajian singkat tentang latar belakang permasalahan, perumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan

BAB II TINJAUAN PUSTAKA

Konsep dasar dan landasan teori yang digunakan sebagai acuan pada penelitian yang dilakukan yaitu; tinjauan pustaka terhadap penelitian-penelitian serupa yang pernah dilakukan sebelumnya, pencarian artikel dan definisi serta teori dasar yang dapat memperkuat penelitian.

BAB III METODOLOGI PENELITIAN

Metodologi yang digunakan dalam penelitian ini, mencakup gambaran umum penelitian, deskripsi data, tahapan dalam melakukan penelitian dan alur metode dalam memproses data.

BAB IV HASIL DAN PEMBAHASAN

Menjelaskan tentang implementasi dari perancangan sistem yang sudah dibuat dan melakukan pengujian terhadap sistem yang sudah dibuat.

BAB V PENUTUP

Berisi kesimpulan dan saran dari hasil penelitian yang sudah dilakukan.