

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Short Message Service (SMS) merupakan sebuah layanan yang banyak diaplikasikan pada sistem komunikasi tanpa kabel, memungkinkan dilakukannya pengiriman pesan dalam bentuk alphanumeric antara terminal pelanggan atau antara terminal pelanggan dengan sistem eksternal seperti SMS, paging, voice mail, dan lain-lain [1]. Pada layanan SMS, Spam sudah menjadi hal yang lumrah dan biasa. Spam pada SMS mengandung pesan yang tidak diinginkan atau tidak diminta oleh pengguna seperti iklan, penipuan, link/tautan berbahaya dan sebagainya. Salah satu faktor yang dapat menyebabkan adanya pesan spam adalah pemberian izin atau permission kepada aplikasi yang tidak jelas ataupun berbahaya sehingga aplikasi tersebut dapat mengakses data pengguna (user).

Berdasarkan penelitian yang dilakukan oleh digital marketing Emarketer memperkirakan pada tahun 2018 adanya peningkatan jumlah pengguna aktif smartphone di Indonesia mencapai lebih dari 100 juta orang. Sehingga menempatkan Indonesia menjadi negara dengan peringkat keempat terbesar di dunia dalam pengguna aktif smartphone setelah Cina, India, dan Amerika. Meningkatnya penggunaan smartphone di Indonesia dapat menimbulkan ancaman privasi data setiap penggunaannya. Salah satu faktor yang menyebabkan terjadinya pelanggaran informasi dan privasi adalah karena para pengguna smartphone memiliki *security awareness* atau kurangnya tingkat kesadaran keamanan dalam menggunakan smartphone yang baik dan aman. Sehingga dengan adanya faktor tersebut dapat menimbulkan indikasi atau peluang terjadinya tindak kejahatan oleh pihak yang tidak bertanggung jawab, salah satunya melalui pesan spam [2].

Pengguna SMS cenderung memiliki kerentanan terhadap penyalahgunaan spam yang berujung pada berbagai kasus *fraud* (penipuan) dan *scam*. Menurut laporan dari National Fraud and Cyber Crime Reporting

Centre (NFCCRC), jumlah kasus penipuan melalui layanan SMS di Inggris meningkat sebesar 700% dari tahun 2016 hingga 2020. Persentase kasus pada layanan SMS lebih banyak dibandingkan dengan kasus pada platform lain seperti pada email yang meningkat sebesar 212% dan pada layanan pesan social media sebesar 43% pada kurun waktu yang sama. [16]

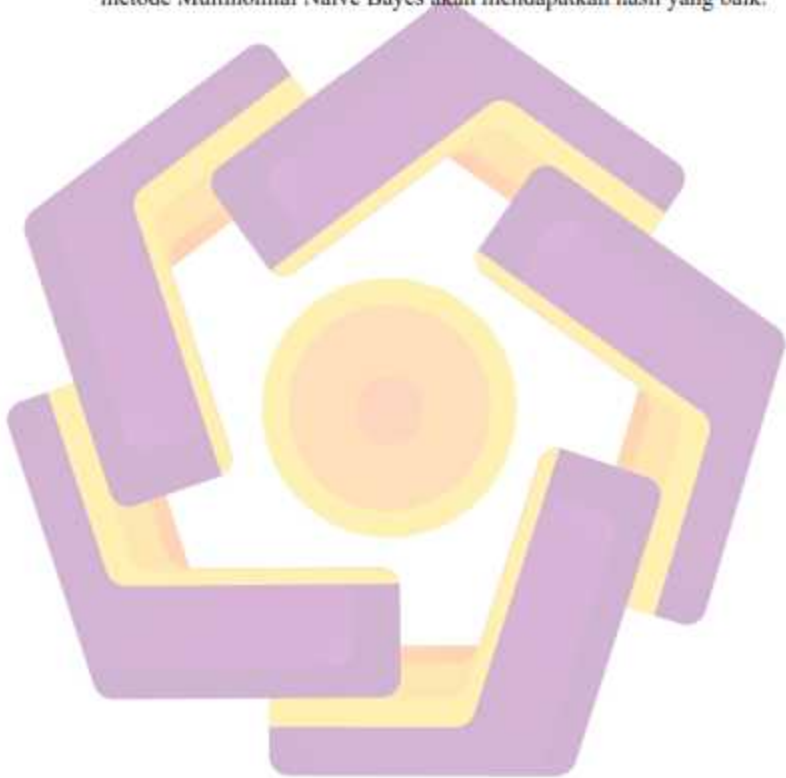
Untuk meminimalisir ketidaknyamanan dan tindak kejahatan yang disebabkan oleh pesan spam, maka perlu dilakukan pengklasifikasian SMS spam atau SMS filtering dengan memfilter SMS spam menggunakan suatu metode untuk memperoleh keakurasian hasil prediksi yang optimal [2].

Berdasarkan hal tersebut, filter spam SMS menjadi suatu hal yang penting di masa sekarang. Ada beberapa cara atau metode dalam filterisasi SMS dengan teknologi komputasi yang ada sekarang seperti metode Naïve Bayes, SVM, Decision Tree, ID3, TAN dll. Fokus pada penelitian ini adalah pada metode Naïve Bayes. Naïve bayes disebut juga multinomial naïve bayes merupakan metode klasifikasi yang memanfaatkan metode probabilitas dan statistik. Metode ini merupakan model penyederhanaan dari algoritma bayes yang cocok dalam pengklasifikasian teks atau dokumen. Dalam klasifikasi menggunakan naïve bayes, nilai kategori dari suatu dokumen akan ditentukan berdasarkan fitur/kata yang muncul dalam dokumen yang akan diklasifikasikan [3].

Spam filtering adalah suatu proses yang memisahkan SMS berdasarkan kategorinya baik spam maupun bukan spam atau biasa disebut ham. Dalam pengklasifikasian SMS spam, dibutuhkan suatu sistem cerdas yang dapat memilah-milah/mengklasifikasikan SMS spam dan bukan spam (ham) secara baik dan benar. Pada proses pengklasifikasian spam SMS tersebut dapat diterapkan beberapa metode/algoritma yang dapat membantu sistem dalam melakukan pengklasifikasian. Pada penelitian ini akan digunakan metode klasifikasi Naïve Bayes. Salah satu alasan penggunaan metode Naïve Bayes ini, adalah karena sifatnya yang *historical* (dapat mengingat kejadian sebelumnya). Hal ini sangat cocok untuk spam filtering, mengingat jenis SMS spam yang bermacam-macam dilihat dari isi SMS dan

headernya [4].

Dengan penelitian ini, diharapkan dapat mencegah atau meminimalisir penyalahgunaan dan dampak negatif dari adanya pesan spam pada platform atau layanan SMS. Berdasarkan kesimpulan yang dipaparkan pada penelitian sebelumnya diharapkan dalam penelitian ini penggunaan metode Multinomial Naïve Bayes akan mendapatkan hasil yang baik.



## 1.2 Rumusan Masalah

1. Bagaimana penerapan algoritma Multinomial Naïve Bayes dalam melakukan klasifikasi untuk membedakan suatu teks berisi pesan spam pada layanan SMS.
2. Bagaimana nilai akurasi dari algoritma Naïve bayes dalam menentukan klasifikasi pesan spam dan ham pada layanan SMS.

## 1.3 Batasan Masalah

Beberapa batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Dataset yang digunakan merupakan dataset umum yang ada pada website UCI Machine Learning.
2. Filter pesan spam layanan SMS pada penelitian kali ini hanya berfokus pada metode Multinomial Naïve Bayes.
3. Proses klasifikasi teks SMS menggunakan algoritma Multinomial Naïve Bayes yang menggunakan bahasa pemrograman Python serta tools IDE Google Colab.
4. Hasil penyaringan data hanya diklasifikasikan kedalam 2 kelas, yaitu pesan spam dan bukan spam (ham).

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

1. Mengetahui apakah metode Multinomial Naïve Bayes dapat melakukan proses filterisasi pesan spam pada layanan SMS.
2. Mencari akurasi yang dihasilkan dari proses filterisasi metode Multinomial Naïve Bayes pada layanan SMS.

## 1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan bermanfaat untuk publik yang merupakan pengguna telepon genggam dan layanan SMS sehingga dapat mencegah maupun meminimalisir kerugian dan dampak negatif dari adanya pesan spam yang beredar pada layanan SMS. Penelitian juga diharapkan dapat menjadi sumber referensi atau dasar pengetahuan bagi penelitian

lanjut yang berkaitan dengan klasifikasi teks pada layanan SMS maupun layanan pengiriman pesan pada platform lain.

## **1.6 Sistematika Penulisan**

Pada bagian ini dituliskan urutan-urutan dan sistematika penulisan yang dilakukan. Berikut ringkasan mengenai masing-masing bab.

### **BAB I PENDAHULUAN**

Bab ini merupakan pengantar terhadap masalah yang akan dibahas, yang terdiri dari latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan, fungsi, metode penelitian, dan sistematika penulisan

### **BAB II LANDASAN TEORI**

Bab ini menguraikan dasar teori, yang dimulai dengan tinjauan puskaka, pengertian dari istilah-istilah yang berkaitan dengan program dan software-software yang akan digunakan dalam pembuatan media penyaringan.

### **BAB III METODE PENELITIAN**

Bab ini menguraikan tentang alat dan bahan yang digunakan pada penelitian, serta alur penelitian.

### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini menjelaskan tentang tahapan dalam pembuatan media penyaringan hingga pengujian media penyaringan SMS berbasis metode Multinomial Naïve Bayes.

### **BAB V PENUTUP**

Bab ini merupakan penutup yang berisi kesimpulan dan saran yang dirangkum selama penelitian.