

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Instagram merupakan salah satu media sosial yang sudah sangat populer di kalangan masyarakat. Berdasarkan survei yang telah dilakukan *We Are Social*, Instagram menduduki peringkat ke-6 yang memiliki pengguna aktif terbanyak [1]. Hampir semua lapisan rakyat dari kalangan manapun memiliki Instagram. Namun, karena media sosial merupakan ruang publik yang mana semua orang dapat mengutarakan pendapatnya di sana, tidak jarang Instagram dijadikan ajang berkomentar yang tidak menyenangkan bahkan dapat tergolong dalam bentuk pelecehan.

Pelecehan seksual yang terjadi di masa kini sudah bukan lagi hanya melalui fisik maupun lisan, namun secara tulisan pun bisa tergolong pelecehan seksual. Pelecehan atau *harrasment* dalam media *online* merupakan salah satu tindakan *cyberbullying* [2]. Berdasarkan data yang diperoleh dari survei *Ditch The Label*, Instagram menduduki peringkat pertama dalam *cyberbullying* dengan presentase sebanyak 42%. [3]

Beragamnya pengguna Instagram, beragam pula tujuannya. Tokoh publik misalnya, mereka memerlukan Instagram salah satunya sebagai media untuk lebih dekat dengan penggemar dan masyarakat, sehingga tidak heran apabila kebanyakan tokoh publik menjadi pengguna 'wajib' di Instagram.

Sebagai orang yang dipandang dalam masyarakat luas, setiap tingkah laku atau setiap postingan bisa jadi memberikan banyak pro dan kontra dalam

masyarakat. Begitu pula mereka harus bersiap menerima komentar apapun yang diajukan bisa positif ataupun negatif.

Reaksi yang ditimbulkan dari postingan yang dibagikan tentu saja sangat beragam dan tidak semuanya positif. Tidak jarang para tokoh publik diserang pengguna yang memberikan respon negatif, seperti dengan melakukan bullying, komentar-komentar asusila yang bisa jadi termasuk dalam pelecehan seksual.

Tindakan pelecehan seksual kebanyakan menyerang perempuan tapi laki-laki pun tak luput dari pelecehan seksual. Misalnya saja salah satu kasus yang terjadi pada salah satu atlit bulu tangkis Indonesia, Jonatan Christie aksi selebrasinya menjadi objek yang menimbulkan banyak komentar dari para pengguna Instagram yang mengutarakan kata-kata seperti "rahim anget", "ovarium meledak" ataupun kata-kata lainnya yang bisa jadi menjurus ke arah pelecehan seksual.

Analisis sentimen sangat diperlukan untuk menyaring komentar-komentar dalam media sosial. Analisis sentimen pada komentar dilakukan untuk mengklasifikasikan komentar yang bersifat positif dan komentar yang bersifat negatif sehingga dapat terdeteksi manakah komentar yang merujuk ke pelecehan seksual mana yang tidak [4]. Dari analisis tersebut diharapkan dapat dilakukan tindakan pencegahan baik bagi korban maupun pelaku dan untuk menertibkan pengguna dalam menggunakan media sosial.

Analisis sentimen dapat dilakukan dengan menggunakan metode untuk klasifikasi. Salah satu metode untuk klasifikasi yang dapat digunakan adalah *Naïve Bayes Classifier*. Berdasarkan beberapa penelitian sebelumnya yang relevan dengan topik ini memberikan hasil akurasi sebesar 80% untuk mengukur data

sebanyak 400 data dari Twitter dengan pengujian sebanyak tiga kali dengan besaran *testing* 10%, 20%, dan 30% [5]. Adapun penelitian lain yang serupa memberikan rata-rata nilai akurasi sebesar 94.78%, presisi 98.86% dan *recall* sebesar 90.98% dari 13 kali percobaan dengan 10 data *training* dan 3 data uji [6].

Dari referensi yang ada, dapat ditarik kesimpulan bahwa metode *Naïve Bayes Classifier* merupakan salah satu metode yang memiliki akurasi tinggi yang dapat digunakan peneliti dalam melakukan analisis. Selain itu, adapun kelebihan lain yang dimiliki oleh Algoritma *Naïve Bayes Classifier* yaitu kecepatan yang tinggi, mudah dipahami, pengkodeannya sederhana, lebih cepat dalam penghitungan, menangani kuantitatif dan data diskrit, dan lain-lain [7].

Berdasarkan uraian di atas, maka penulis akan melakukan penelitian sebagai skripsi dengan judul **"Implementasi *Naïve Bayes Classifier* untuk Mendeteksi Komentar Pelecehan Seksual terhadap Tokoh Publik di Instagram"**.

## 1.2 Rumusan Masalah

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

1. Bagaimana penerapan algoritma *Naïve Bayes Classifier* dalam mendeteksi pelecehan seksual di Instagram pada tokoh publik?
2. Bagaimana tingkat akurasi sistem dalam menerapkan algoritma *Naïve Bayes Classifier* untuk mendeteksi pelecehan seksual di Instagram pada tokoh publik?

### 1.3 Batasan Masalah

Agar penelitian ini dapat terfokus maka diperlukan adanya batasan-batasan untuk membatasi ruang lingkup dari penelitian ini. Adapun batasan-batasan dari penelitian ini adalah sebagai berikut :

1. Menggunakan metode klasifikasi yaitu metode *Naïve Bayes Classifier* untuk mengklasifikasikan komentar yang bersifat pelecehan seksual (*harassment*) dan bukan pelecehan seksual (*not harassment*).
2. Objek yang diteliti merupakan postingan dari tokoh publik di Indonesia.
3. Komentar yang dipilih merupakan komentar berbahasa Indonesia.
4. Pembuatan sistem ini berbasis *website*.
5. Data *training* penelitian ini menggunakan 15 postingan dari 15 tokoh publik di Indonesia yang diambil pada tahun 2018-2019.
6. Pengolahan dan pelabelan data *harassment* dilakukan secara manual berdasarkan studi literatur dari beberapa sumber.
7. Sistem yang dibangun berbasis *website* menggunakan bahasa pemrograman Javascript.

### 1.4 Maksud dan Tujuan Penelitian

Maksud dan tujuan dari penelitian ini adalah sebagai berikut :

1. Membangun sistem yang dapat mengklasifikasikan komentar terhadap tokoh publik yang merujuk ke arah pelecehan seksual atau pun tidak menggunakan *Naïve Bayes Classifier*.
2. Memberikan evaluasi dari hasil klasifikasi dari sistem yang diusulkan.

## **1.5 Metodologi Penelitian**

Untuk mendukung penelitian ini, penulis menggunakan beberapa metodologi penelitian, di antaranya adalah:

### **1.5.1 Metode Pengumpulan Data**

Metode pengumpulan data merupakan metode yang diperlukan untuk memperoleh informasi yang dibutuhkan untuk menunjang penelitian. Adapun metode pengumpulan data yang akan digunakan dalam penelitian ini adalah metode studi pustaka. Studi pustaka atau studi kepustakaan adalah proses menghimpun informasi yang relevan topik atau permasalahan dari buku-buku, karya ilmiah, jurnal-jurnal, skripsi maupun situs internet yang dapat membantu penelitian. Dalam hal ini penulis mencari informasi mengenai *Naïve Bayes Classifier*, Instagram, *Confusion Matrix*, teori pelecehan seksual, tokoh publik, dan informasi lain yang relevan.

#### **1.5.1.1 Metode Observasi**

Observasi merupakan bagian dari teknik pengumpulan data, dimana peneliti melakukan pengamatan secara langsung ke objek penelitian untuk mengamati dan mengumpulkan data yang dibutuhkan. Adapun penulis akan melakukan observasi terhadap postingan-postingan tokoh publik di Instagram yang relevan dengan penelitian.

### **1.5.2 Metode Analisis**

Metode analisis merupakan tahapan dalam menganalisis informasi yang diperoleh dalam proses pengumpulan data untuk membangun model yang akan dibuat. Dalam metode ini akan dideskripsikan mengenai implementasi metode yang digunakan dalam model penelitian.

### 1.5.3 Metode Perancangan

Pada penelitian ini dalam membangun model klasifikasi komentar pelecehan seksual, penulis akan menggunakan *flowchart* dan *Data Flow Diagram* (DFD).

*Flowchart* merupakan penguraian dari model utama ke sub-model dengan tujuan untuk mengidentifikasi permasalahan-permasalahan yang ada dan kebutuhan-kebutuhan yang diperlukan agar dapat dievaluasi dan diusulkan menjadi ke dalam model yang lebih baik. Sedangkan DFD akan menerangkan aliran data dari program penelitian ini.

### 1.5.4 Metode Pengujian

Pengujian yang dilakukan pada penelitian ini menggunakan *Confusion Matrix* dan *K-Fold Cross Validation*. Kedua metode ini merupakan metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi.

## 1.6 Sistematika Penulisan

Penelitian ini dituliskan dalam suatu sistematika penulisan sebagai berikut:

### **BAB I PENDAHULUAN**

Pada bab ini berisi mengenai latar belakang penelitian, rumusan penelitian, batasan penelitian, maksud dan tujuan penelitian, metode penelitian yang digunakan serta sistematika penulisan penelitian.

### **BAB II LANDASAN TEORI**

Pada bab ini berisi kajian pustaka yang berisi penelitian-penelitian lain yang relevan serta teori-teori yang dijadikan dasar dalam penelitian ini.

### **BAB III ANALISIS DAN PERANCANGAN**

Pada bab ini memberikan penjelasan mengenai analisis kebutuhan dan perancangan sistem yang akan diusulkan dalam penelitian ini.

#### **BAB IV IMPLEMENTASI DAN PEMBAHASAN**

Pada bab ini berisi mengenai implementasi rancangan sistem yang akan dibuat serta akan dijelaskan pula mengenai pengujian dan evaluasi dari sistem yang akan dibuat.

#### **BAB V PENUTUP**

Pada bab ini berisi kesimpulan dan saran dari penulis mengenai penelitian yang telah dilakukan.

#### **DAFTAR PUSTAKA**

Berisi sumber-sumber yang digunakan penulis dalam melakukan penelitian.

