

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi informasi (*internet*) khususnya media sosial mempunyai dua sisi bagaikan mata pisau. Pada satu sisi berdampak positif, yaitu mempunyai manfaat untuk mempermudah manusia dalam bertukar informasi menjadi lebih cepat, berinteraksi sesama pengguna, dan berbagai aktivitas lain yang memungkinkan manusia menciptakan karya mereka sesuai dengan kreatifitasnya. Pada sisi lain yang tidak bisa dihindarkan adalah dampak negatif dari teknologi informasi (*internet*) khususnya sosial media, yaitu tindakan yang melanggar kebijakan dalam menggunakan sosial media dan kejahatan dunia maya (*cyber crime*) [1]. Salah satu penyalahgunaan tersebut adalah tindakan prostitusi yang dilakukan melalui jaringan internet atau disebut prostitusi online. Menurut Kamus Besar Bahasa Indonesia (*KBBI*) prostitusi adalah pertukaran hubungan seksual dengan mata uang atau hadiah sebagai suatu transaksi perdagangan; pelacuran. Sedangkan prostitusi online merupakan kegiatan prostitusi atau suatu kegiatan yang menjadikan seseorang sebagai objek untuk diperdagangkan melalui media elektronik atau online dalam hal ini media sosial.

Saat ini sosial media menjadi bagian dari kehidupan sosial manusia, salah satu sosial media yang populer di Indonesia adalah *Twitter*. Pada laporan finansial *Twitter* kuartal ke-3 tahun 2019, pengguna aktif harian di platform *Twitter* tercatat mengalami kenaikan hingga 17 persen, ke angka 145 juta pengguna. Menurut Dwi Ardiansah selaku *Country Industry Head Twitter*

*Indonesia* pada Oktober 2019, *Indonesia* diklaim menjadi salah satu negara yang pertumbuhan pengguna aktif harian *Twitter*-nya paling besar [2]. Dengan pertumbuhan pengguna aktif yang terus bertambah, berbanding lurus dengan pemanfaatan pengguna dalam kebebasan bersosial tersebut. Masalahnya adalah banyak yang memanfaatkan untuk hal-hal negatif salah satunya adalah prostitusi. Bahayanya jika ada anak yang dibawah umur mengakses hal tersebut di sosial media *Twitter* karena kebebasan mereka dalam menjelajah dan memakai sosial media.

Di *Indonesia*, para pelaku yang memasarkan jasa prostitusi online pada *Twitter* dengan melakukan *tweet* di akunnya menggunakan kata-kata dan memiliki tagar sendiri, biasanya mengacu pada daerah dimana domisili mereka tinggal atau tempat mereka menawarkan jasanya [3]. Misalnya di lokasi *Yogyakarta*, mereka memajangkan kiriman *tweet* dengan tagar *#bojogja*, *#openbojogja*, *#angellojogja* dan masih banyak lagi. Pihak kepolisian serta Kementerian Komunikasi dan Informatika (Kominfo) sebenarnya telah melakukan banyak upaya untuk meredam aksi prostitusi *online* dengan melakukan pemblokiran akun. Pihak *Twitter* pun telah melakukan upaya pencegahan dengan fitur *report*, tetapi upaya ini kurang efektif mengingat budaya masyarakat yang terkesan acuh dengan akun-akun semacam ini.

Dengan fenomena seperti ini, dibutuhkan suatu metode untuk mengklasifikasi *tweet* dan akun *Twitter* untuk mengetahui dan mendeteksi *tweet* dan akun-akun yang masuk dalam kategori jasa prostitusi ini guna membantu menangani masalah tersebut. Beberapa metode yang digunakan untuk klasifikasi data antara lain *Support Vector Machine (SVM)*, *Naive Bayes*, *K-Nearest*

*Negihbor (KNN), dan Decission Tree*. Dari setiap metode memiliki kelebihan dan kekurangan masing-masing. Beberapa penelitian tentang klasifikasi data twitter seperti telah banyak dilakukan oleh peneliti sebelumnya dengan metode yang berbeda-beda, contohnya penelitian yang dilakukan oleh Yufis Azhar dengan judul *Klasifikasi Akun Prostitusi Berdasarkan Skoring Tweet* [4], menggunakan metode *K-Nearest Neighbor (KNN)* dan *TF.IDF* untuk pembobotan *tweet* dapat menghasilkan klasifikasi akun dengan tingkat akurasi cukup tinggi tergantung dengan jumlah data latih yang digunakan.

Pemilihan metode *Naive Bayes* sebagai metode yang digunakan pada penelitian ini dikarenakan metode *Naive Bayes* sangat sesuai untuk melakukan klasifikasi data *Twitter* dan klasifikasi data teks dalam jumlah besar. Selain itu, menurut (Akhmad Pandu Wijaya & Diki Wardhani, 2019) melalui penelitiannya dengan judul *Analisa Sentimen dan Klasifikasi Komentar Positif Pada Twitter dengan Naive Bayes Classification* [5], menerangkan bahwa *Naive Bayes* memiliki performa yang baik dalam melakukan klasifikasi data tanpa ada penambahan metode lain.

Berdasarkan latar belakang yang telah diuraikan diatas, maka penulis mencoba melakukan penelitian untuk menganalisa tindakan tindakan prostitusi online pada *Twitter* menggunakan metode *Naive Bayes*. Sehingga, dapat diketahui hasil dari klasifikasi data-data tersebut dan dapat diukur tingkat akurasinya pada dataset tindakan prostitusi online di *Twitter*.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka masalah yang akan diselesaikan pada penelitian ini adalah :

1. Bagaimana mengimplementasikan metode *Naive Bayes* untuk melakukan klasifikasi *tweet* atas tindakan prostitusi *online* pada *Twitter*.
2. Berapa tingkat akurasi metode *Naive Bayes* dalam melakukan klasifikasi *tweet*.

## 1.3 Batasan Masalah

Dalam penulisan penelitian ini penulis akan mengimplementasikan algoritma *Naive Bayes* kedalam program. Untuk itu agar mendapatkan arah pembahasan yang lebih baik sehingga tujuan penulisan penelitian dapat tercapai, maka penulis membatasi ruang lingkup permasalahan yang ada yaitu sebagai berikut :

1. Data yang di proses hanya berupa teks (tidak termasuk file gambar ataupun video).
2. Data yang dianalisis diambil di wilayah Yogyakarta dengan tagar khusus *#bojogja*, *#angeljogja* dan *#openbojogja*.
3. Data *tweet* hanya menggunakan Bahasa Indonesia atau kata yang sudah ditentukan sebelumnya.
4. Metode klasifikasi yang digunakan menggunakan *Naive Bayes* dan tidak membandingkan dengan metode lain.

5. Data yang dianalisis tidak melalui tahap analisis *typo correction* ataupun singkatan kata.

#### 1.4 Maksud dan Tujuan Penelitian

Maksud dari penelitian ini adalah mengimplementasikan metode *Naive Bayes* sebagai sistem klasifikasi *tweet* yang termasuk ke dalam tindakan prostitusi pada *Twitter*.

Tujuan dari penelitian ini adalah membantu memecahkan permasalahan penyalahgunaan *Twitter* sebagai media untuk kegiatan prostitusi *online* dan kegiatan pelacuran.

#### 1.5 Manfaat Penelitian

Adapun manfaat yang bisa diperoleh dari penelitian ini, diantaranya :

1. Bagi Pengguna Sistem

Penelitian ini dapat dijadikan sebagai salah satu alternatif metode dalam mengidentifikasi tindakan prostitusi *online* pada *Twitter*.

2. Bagi Peneliti

Penelitian ini digunakan sebagai tugas akhir atau skripsi sebagai syarat kelulusan di program sarjana, Universitas Amikom Yogyakarta dan menambah pengetahuan peneliti mengenai klasifikasi data menggunakan metode *Naive Bayes*.

3. Bagi Universitas Amikom

Penelitian ini diharapkan dapat menjadi referensi dari penelitian selanjutnya, terutama bagi mahasiswa Universitas Amikom Yogyakarta

yang ingin mengadakan penelitian yang menggunakan metode *Naive Bayes*.

## **1.6 Metode Penelitian**

Pada penelitian ini penulis menggunakan beberapa metode penelitian sebagai berikut :

### **1.6.1 Metode Pengumpulan Data**

Menggunakan metode studi literatur dengan mengumpulkan materi dan mempelajari teori-teori mengenai *data mining*, *text processing*, *Naive Bayes*. Sumber bacaan berupa *ebook*, *jurnal*, dan *artikel* dari internet guna menunjang penelitian. Mengumpulkan data dari *Twitter* tentang tindakan prostitusi online dengan kata kunci tertentu yang meliputi kata yang sering digunakan, dan tagar yang digunakan.

### **1.6.2 Metode Analisis**

Data maupun informasi yang didapatkan kemudian dipelajari dan dianalisa. Agar model yang direncanakan diketahui maka diperlukan implementasi dari model ke bentuk program komputer. Analisis disini merupakan analisis data yang berasal dari data yang didapat dari *Twitter* lalu diolah menggunakan algoritma *Naive Bayes*. Lalu juga dilakukan analisis sistem yang mencakup analisis fungsional dan non fungsional.

### **1.6.3 Metode Perancangan**

Metode perancangan berisi paparan deskriptif langkah-langkah dalam proses perancangan program aplikasi. Langkah-langkah ini meliputi latar belakang atau ide perancangan, identifikasi permasalahan, tujuan perancangan

dan konsep perancangan. Perancangan sistem klasifikasi tweet dengan metode Naïve Bayes dimulai dengan merubah setiap kata pada tweet kedalam kata dasar, pengkoreskian kesalahan penulisan, menghilangkan simbol dan kemudian dilakukan pembobotan setiap kata.

#### **1.6.4 Metode Implementasi**

Mengumpulkan data uji dan data latih dari *Twitter* kemudian mengolahnya sehingga menghasilkan pengembangan sistem dengan materi yang sudah dirancang.

#### **1.6.5 Metode Testing**

Pengujian adalah proses yang bertujuan untuk memastikan semua fungsi sistem berjalan sesuai harapan. Untuk melakukan pengujian akurasi dalam klasifikasi data, digunakan metode pengujian *Confusion Matrix*, *Cross Validation*, dan menguji perbandingan *data training*.

### **1.7 Sistematika Penulisan**

Untuk memudahkan dalam mengikuti seluruh uraian dan pembahasan pada penelitian ini, maka penulisan penelitian ini dilakukan dengan sistematika sebagai berikut:

#### **BAB I PENDAHULUAN**

Pada bab ini penulis menerangkan tentang latar belakang penelitian, perumusan masalah, pembatasan masalah, maksud tujuan penelitian, metode penelitian, serta sistematika penulisan.

#### **BAB II LANDASAN TEORI**

Pada bab ini akan diuraikan tentang teori-teori pendukung yang berkaitan dengan tema dan akan dijadikan dasar penelitian ini. Teori yang akan diangkat yaitu mengenai prostitusi *online* dengan algoritma *Naives Bayes*.

### BAB III ANALISIS DAN PERANCANGAN

Pada bab ini akan dijelaskan tentang tujuan perancangan sistem, analisa kebutuhan sistem, perancangan perangkat lunak dan juga tahapan dalam mengimplementasikan metode.

### BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini akan membahas mengenai implementasi sistem prostitusi *online* dengan menggunakan algoritma *Naives Bayes*. Hasil analisis menggunakan metode tersebut akan dibahas pada bab ini.

### BAB V PENUTUP

Pada bab ini berisi kesimpulan dari skripsi yang dibuat dan menjelaskan saran saran penulis kepada pembaca agar hasil dari analisis ini dengan menggunakan algoritma *Naives Bayes* dapat bermanfaat dalam pengembangan lebih lanjut.

### DAFTAR PUSTAKA

Berisi sumber bacaan yang digunakan penulis sebagai bahan penelitian.