

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pesatnya perkembangan teknologi di era sekarang banyak membantu segala kegiatan manusia. Teknologi berperan penting dalam kehidupan manusia, salah satunya dalam bidang komunikasi. Dari peran penting teknologi pada bidang komunikasi ini, melahirkan wadah baru yang kita kenal sebagai media sosial. Menurut[1], media sosial adalah suatu sarana media *online* yang digunakan oleh manusia agar dapat dengan mudah berpartisipasi, berbagi, menciptakan isi seperti blog, jejaring sosial, *wiki*, forum, dan dunia virtual.

Dari perkembangan teknologi tersebut, masih banyak oknum yang menyalahgunakan perkembangan tersebut. Tindakan yang paling sering terjadi adalah dalam bentuk intimidasi. Menurut KBBI, intimidasi bermakna tindakan menakut-nakuti, gertakan, dan ancaman. Tindakan seperti ini dilakukan pada jejaring sosial dan disebut sebagai *cyberbullying*. Berdasarkan laporan yang dilakukan oleh *Microsoft*, Indonesia berada di urutan ke-29 dari 32 negara yang disurvei berdasarkan kesopanan pada media sosial[2].

Maka dari uraian tersebut, peneliti ingin melakukan pengklasifikasian komentar yang diinputkan oleh netizen apakah termasuk komentar *bullying* atau *non-bullying*. Peneliti menggunakan algoritma Naïve Bayes Classifier model Multinomial dan Bernoulli untuk menguji kinerja dari model tersebut karena menurut[3] algoritma Naïve Bayes hanya membutuhkan jumlah data pelatihan yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian.

Untuk meningkatkan performa dari model klasifikasi yang dibuat maka peneliti menggunakan 2 metode *Feature Extraction* dan membandingkan performanya pada algoritma yang digunakan. Metode *Feature Extraction* yang digunakan adalah *Bag of Words* dan TF-IDF (*Term Frequency - Inverse Document*

Frequency). *Feature Extraction* diperlukan karena komputer hanya mampu memproses data yang numerik sehingga data berupa teks perlu direpresentasikan menjadi numerik menggunakan metode *Bag of Words* dan TF-IDF.[4]

1.2 Rumusan Masalah

Adapun yang menjadi rumusan masalah adalah sebagai berikut :

1. Bagaimana perbandingan kinerja antara algoritma Naïve Bayes Classifier model Multinomial dan Bernoulli dalam mengklasifikasikan data komentar *cyberbullying*?
2. Bagaimana perbandingan penggunaan *Feature Extraction* dengan metode *Bag of Words* dan TF-IDF dalam peningkatan performa model yang digunakan?

1.3 Batasan Masalah

Batasan masalah dalam penulisan skripsi adalah :

1. *Dataset* yang digunakan adalah *dataset* dari *website Kaggle* dengan 650 jumlah data dan sudah memiliki label.
2. Data yang diteliti adalah *dataset* komentar media sosial *instagram* yang berbahasa Indonesia.
3. Penelitian yang dilakukan hanya berfokus pada komentar yang termasuk *cyberbullying* atau tidak.
4. Pembagian label kelas hanya menjadi 2 (dua) label yaitu *Bullying* dan *Non-Bullying*.

1.4 Tujuan Penelitian

Adapun tujuan yang diinginkan peneliti pada penelitian ini sebagai berikut :

1. Membandingkan kinerja algoritma Naïve Bayes model Multinomial dan Bernoulli.

2. Membuat program berbasis *web html*.
3. Untuk mengetahui komentar yang diinputkan termasuk komentar *cyberbullying* atau tidak.
4. Membandingkan penggunaan metode *Feature Extraction* dalam peningkatan performa model klasifikasi yang dibuat.

1.5 Manfaat Penelitian

Adapun beberapa manfaat yang diharapkan peneliti dari penelitian ini sebagai berikut :

1. Bagi Peneliti :
 - a. Agar memahami cara kerja Algoritma Naive Bayes Classifier dalam melakukan klasifikasi data berupa teks.
 - b. Menambah ilmu dan wawasan peneliti
2. Bagi Masyarakat Umum/Pihak tertentu :
 - a. Agar mengetahui komentar yang diketik oleh para *netizen* (pengguna internet) mengandung *cyberbullying* atau tidak.
 - b. Dapat menjadi acuan untuk mengurangi tingkat *cyberbullying* yang terjadi.
3. Bagi Universitas :
 - a. Sebagai pedoman untuk penelitian selanjutnya.

1.6 Sistematika Penulisan

Adapun sistem penulisan pada skripsi ini sebagai berikut :

BAB I PENDAHULUAN

Berisi tentang latar belakang permasalahan yang diangkat, perumusan masalah yang akan diteliti, tujuan dilakukannya penelitian, manfaat dari penelitian

yang dilakukan dan sistematika dari penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Berisi tentang konsep dasar dan landasan teori dari penelitian sebelumnya yang berkaitan atau relevan dengan topik penelitian dan digunakan sebagai acuan pada penelitian yang dilakukan.

BAB III METODE PENELITIAN

Berisi penjelasan tentang metode yang digunakan, gambaran singkat tentang penelitian, alur dan tahapan penelitian, alat dan bahan yang diperlukan pada penelitian dan menganalisis apakah penelitian sudah sesuai dengan metode yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Berisi penjelasan hasil dari penelitian yang telah dilakukan, penyajian hasil penelitian dan menjelaskan tentang implementasi metode yang digunakan.

BAB V PENUTUP

Berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya

