

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ujaran kebencian akhir-akhir ini banyak menyita perhatian. Salah satu kasus ujaran kebencian berbasis media sosial yang populer di Indonesia adalah kasus ujaran kebencian yang dilakukan oleh musisi Ahmad Dhani. Ahmad Dhani membuat cuitan di Twitter, pada Maret 2017 lalu. Cuitan tersebut secara tidak langsung merujuk pada Gubernur DKI nonaktif saat itu, Basuki Tjahaja Purnama. Kasus ujaran kebencian ini telah ditangani oleh kepolisian setelah pihak Basuki Tjahaja Purnama melapor ke Polisi.

Ujaran kebencian sendiri didefinisikan sebagai komunikasi yang bertujuan meremehkan orang, kelompok, atau golongan berdasarkan suku, agama, ras, etnik, golongan, kewarganegaraan, dan karakteristik lain [1]. Bertambahnya pengguna internet setiap tahunnya, berdampak pada meningkatnya jumlah ujaran kebencian yang tersebar di media sosial [2]. Dengan banyaknya ujaran kebencian yang tersebar di internet, tentu saja membuat baik korban maupun pengguna media sosial yang melihat merasa tidak nyaman.

Hal yang membuat klasifikasi ujaran kebencian sangat sulit, karena tidak ada standar yang benar-benar baku untuk ujaran kebencian [3]. Beberapa merasa komentar/cuitan seseorang di media sosial sangat menyakitkan, namun mungkin bagi orang lain, hal tersebut bukan masalah. Tidak banyak korban ujaran kebencian yang melapor, baik karena takut atau merasa tidak peduli, membuat perilaku ujaran kebencian akan selalu tumbuh subur di Indonesia.

Indonesia sendiri, sudah memiliki UU ITE (Undang-Undang Informasi Transaksi Elektronik) yang dapat digunakan untuk menjerat para pelaku ujaran kebencian ke meja hijau. Salah satu cara lain untuk menghentikan ujaran kebencian adalah dengan melakukan penyaringan komentar atau cuitan. Hal yang paling

mudah untuk membatasi ujaran kebencian adalah melakukan pemblokiran cuitan untuk setiap kata/frasa yang dianggap sebagai bagian dari ujaran kebencian. Namun, karena tidak semua komentar maupun cuitan yang menggunakan kata/frasa tersebut. Sebagai contoh, menggunakan kata anjing dalam konteks hewan peliharaan akan sangat berbeda bila kata anjing digunakan untuk menyatakan umpatan/kebencian.

Pembelajaran mesin diperlukan untuk membantu media sosial dalam melacak ujaran kebencian, bahkan melakukan upaya preventif, yaitu menghindari ujaran kebencian tersebut tampil di media sosial. Alasan digunakannya pembelajaran mesin, karena melalui pembelajaran mesin, sistem dapat melakukan analisis untuk mengetahui apakah komentar/cuitan seseorang mengandung ujaran kebencian. Dataset ujaran kebencian juga diperlukan sehingga mesin dapat mempelajari mana saja yang termasuk ujaran kebencian. Untuk meningkatkan akurasi, maka dilakukan pengujian dengan menggunakan berbagai algoritma pembelajaran mesin, sehingga dapat ditemukan algoritma yang akurat dalam mendeteksi ujaran kebencian.

Model pembelajaran mesin deteksi ujaran kebencian ini dapat digunakan dan dimanfaatkan siapa saja demi menciptakan internet yang bebas dari ujaran kebencian. Model untuk mencegah ujaran kebencian ini akan dibangun menggunakan Bahasa pemrograman Python dan *tools* Jupyter Notebook.

1.2 Rumusan Masalah

Dengan merujuk latar belakang yang terdapat di atas, rumusan masalah yang didapatkan adalah:

1. Bagaimana membuat dataset untuk model pembelajaran mesin deteksi ujaran kebencian.
2. Bagaimana membuat model untuk mengenali ujaran kebencian.
3. Bagaimana mengevaluasi model untuk menemukan algoritma yang paling baik.

1.3 Batasan Masalah

Batasan-batasan dari masalah yang dirumuskan adalah:

1. Dataset merujuk pada cuitan dalam Bahasa Indonesia.
2. Dataset ujaran kebencian yang dibuat diambil dari Twitter.
3. Algoritma pembelajaran mesin yang akan digunakan dalam penelitian ini adalah: Naïve Bayes, Logistic Regression, dan Support Vector Machines.
4. Parameter dalam dataset yang digunakan untuk melakukan deteksi ujaran kebencian adalah cuitan, subjek yang dimaksud, pembuat cuitan, tagar, dan waktu cuitan dibuat.
5. Dataset menggunakan *feature extraction* TF-IDF, Word2Vec dan N-gram yang akan diimplementasikan pada masing-masing algoritma.

1.4 Tujuan Penelitian

Tujuan akhir dari penelitian adalah:

1. Membuat dataset untuk model pembelajaran mesin deteksi ujaran kebencian.
2. Membuat model untuk mengenali ujaran kebencian.
3. Mengevaluasi model untuk menemukan algoritma yang paling baik

1.5 Sistematika Penelitian

Laporan ini disusun secara sistematis dengan urutan sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metodologi penelitian, dan sistematika penulisan laporan.

BAB 2 TINJAUAN PUSTAKA

Bab ini berisi penjelasan dari penelitian-penelitian yang sebelumnya telah dilakukan akan digunakan untuk pemecahan masalah.

BAB 3 LANDASAN TEORI

Bab ini berisi dasar-dasar teori yang digunakan sebagai pedoman dan acuan dalam pemecahan masalah.

BAB 4 METODOLOGI PENELITIAN

Bab ini berisi penjelasan mengenai metode, desain, dan perancangan dari model yang dibuat.

BAB 5 HASIL PENELITIAN DAN PEMBAHASAN

Bab ini berisi penjelasan mengenai pengujian model.

BAB 6 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan mengenai model yang telah dibuat beserta saran-saran yang bermanfaat untuk pengembangan lebih lanjut.