

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pada 25 Agustus 2017, Bareskrim polri menyatakan telah menangkap tiga orang dari kelompok Saracen, yang merupakan sindikat penyedia jasa konten kebencian berdasarkan suku, agama, ras dan antar golongan. Mereka mempunyai 2000 akun media sosial yang kemudian berkembang menjadi 800000 akun, yang digunakan untuk menyebarkan konten kebencian [1].

Salah satunya konten yang dibagikan adalah ujaran kebencian (*hatespeech*), yang mengekspresikan rasa benci terhadap sekelompok orang tertentu, konflik agama dan etnis yang berakibat pencemaran nama baik pada yang bersangkutan. Berdasarkan pasal 27 ayat (3) UU ITE “Setiap orang dengan sengaja dan tanpa hak mendistribusikan atau mentransmisikan atau membuat dapat diaksesnya informasi elektronik atau dokumen elektronik yang bermuatan penghinaan dan/atau pencemaran nama baik” [2].

Dengan tersedianya fitur unggah gambar pada media social (*Facebook*, *Twitter* dan *Instagram*), beberapa oknum telah memanfaatkan fitur tersebut dengan tidak bertanggung jawab dalam penyebaran ujaran kebencian diberbagai media social, karena *image* dinilai sebagai media paling menarik dan mudah diterima oleh masyarakat.

Solusi dari banyaknya penyebaran ujaran kebencian dipublik yaitu dengan membuat sistem deteksi ujaran kebencian berbasis image to text. Pada sistem tersebut menggunakan input berupa image yang mengandung teks dimana nantinya akan dikonversikan menjadi teks, kemudian teks tersebut akan dianalisa, apakah teks tersebut mengandung ujaran kebencian atau tidak dengan menggunakan metode Latent Semantic Analysis (LSA). LSA adalah metode statistic dan matematika yang digunakan untuk mengekstraksi makna kata-kata kontekstual dan untuk mengidentifikasi hubungan antara kata-kata dalam teks. LSA dapat dianggap sebagai metode tanpa pengawasan untuk mengenali *hate* (benci) dalam teks sehingga waktu tidak boleh dihabiskan untuk tugas pemberian anotasi dari kumpulan data pelatihan.

1.2 Rumusan Masalah

Tujuan dari penyusunan Tugas Akhir ini adalah sebagai berikut :

1. Bagaimana mendeteksi sebuah kalimat pada image ?
2. Bagaimana mendeteksi kalimat yang mengandung ujaran kebencian atau tidak mengandung ujaran kebencian pada image ?

1.3 Tujuan

1. Membuat sebuah aplikasi berbasis web yang dapat mendeteksi kalimat ujaran kebencian pada image yang mengandung teks.
2. Menjadikan sebuah alat bantu validasi untuk mendeteksi ujaran kebencian.

1.4 Batasan Masalah

1. Data-data yang digunakan pada penelitian ini adalah data-data yang didapatkan dari hasil data mining akun-akun acak yang ada pada Twitter berdasarkan kata kunci yang sudah ditentukan.
2. Sistem yang akan dibuat dalam bentuk aplikasi berbasis web.
3. Input pada sistem berupa image dengan format JPG, JPEG dan PNG.
4. Image yang terdapat teks didalamnya seperti ucapan atau kalimat.
5. Teks yang dianalisis hanya dikategorikan menjadi 2 yaitu yang mengandung ujaran kebencian atau tidak mengandung ujaran kebencian.
6. Teks yang dianalisis hanya teks berbahasa Indonesia.
7. Teks yang bisa dikenali oleh sistem adalah teks yang penulisannya jelas dengan ukuran yang sudah ditentukan tidak melebihi dari ukuran *font* 36.
8. Beberapa kata yang sulit untuk di *stemming* seperti berikan dan beruang. Pada berikan bisa saja ber-ikan, beri-kan, ber-i-kan, sedangkan pada beruang bisa saja ber-uang atau nama sebuah binatang beruang.
9. *User* melakukan seleksi *crop* pada bagian image yang mengandung teks

1.5 Metode Penelitian

Pada Tugas Akhir ini penulis menerapkan beberapa metode untuk menyelesaikan masalah sebagai berikut :

1. Studi Literatur

Selama penelitian ini berlangsung, penulis mempelajari konsep-konsep dari sumber literature penunjang materi seperti *crawling data Twitter*, *pre-processing* dan metode *Latent Semantic Analysis*. Sumber *literature* itu meliputi buku-buku, jurnal, *paper*, *website*, *blog* dan dokumen yang sudah ada sebelumnya, ini semua bertujuan agar penulis mendapatkan pemahaman yang cukup.

2. Pengumpulan Data

Pengambilan data dilakukan melalui *API Streaming* pada *Twitter*. Hal ini memungkinkan untuk mendapatkan data dari berbagai sumber akun-akun. Data yang diperoleh akan disimpan kedalam *database*. Pengumpulan data dalam penelitian ini dilakukan dengan mengacu pada hasil kata kunci (*query*) yang sudah ditentukan.

3. Perancangan Sistem dan Realisasi Sistem

Pada tahap ini penulis membuat rencana perancangan sistem yang akan dibuat, mulai dari analisis teori, desain, implementasi dan pengujian.

4. Pengujian dan Analisa

Pada tahap ini akan direalisasikan rancangan sistem yang telah dibuat sebelumnya, mengimplementasikan teks pada image yang mengandung ujaran kebencian atau tidak mengandung ujaran kebencian.

1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam Tugas Akhir ini adalah ;

BAB I PENDAHULUAN

Bab ini membahas mengenai : latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian.

BAB II LANDASAN TEORI

Bab ini membahas mengenai dasar-dasar teori yang akan digunakan pada penelitian ini untuk memecahkan masalah yang diambil dari berbagai sumber.

BAB III PERANCANGAN SISTEM

Bab ini berisi tentang gambaran umum sistem, spesifikasi dan pengambilan data untuk keperluan sistem, perhitungan LSA.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini membahas mengenai implementasi sistem dan pengujian kinerja aplikasi.

BAB V KESIMPULAN DAN SARAN

Bab ini membahas mengenai kesimpulan yang bias ditarik dari penelitian ini dan memberikan saran-saran yang bias membangun penelitian selanjutnya.