

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini situs *microblogging* telah menjadi alat komunikasi yang sangat populer di kalangan pengguna internet. Jutaan pesan muncul setiap hari di situs *microblogging* seperti Twitter, Tumblr, dan Facebook (Manalu, 2014). Twitter menjadi *microblogging system* yang paling populer tentunya di kalangan pengguna jejaring sosial khususnya yang ingin tampil eksis seperti anak muda, pelajar, masyarakat umum, selebriti, bahkan petinggi negara dapat mem-*posting* sesuatu ke dalamnya.

Twitter adalah layanan jejaring sosial yang memungkinkan pengguna untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter dan biasanya juga dikenal dengan *tweet* yang berarti kicauan para pengguna Twitter. Twitter mengalami pertumbuhan yang pesat hingga pada bulan Januari 2015, terdapat lebih dari 500 juta pengguna yang sudah terdaftar di Twitter yang mana 302 juta diantaranya merupakan pengguna aktif. Berdasarkan data PT. Bakrie Telecom, Indonesia menempati urutan ke lima setelah USA, Brazil, Jepang, dan Inggris sebagai negara pengguna Twitter terbesar di dunia. Twitter sudah menjadi tempat untuk melakukan bertukar pendapat dan pikiran tentang topik atau suatu permasalahan, dan mengungkapkan perasaan atau emosi yang sedang dialami dalam berbentuk teks melalui *tweet* yang mereka *posting*.

Teks merupakan salah satu media yang digunakan untuk berkomunikasi dan menyampaikan informasi. Tidak hanya informasi, teks juga dapat mengekspresikan emosi (Bata, Suyoto & Pranowo, 2015). Menurut Morgan dkk, (1968) menyatakan bahwa emosi itu adalah mengatakan sesuatu tentang apa yang kita rasakan ketika kita sedang emosional, menyebut secara psikologis atau secara badaniah, dasar dari perasaan emosional, emosi berpengaruh terhadap persepsi, pikiran, dan perilaku, menjelaskan dorongan atau motivasional, perlengkapan dari emosi-emosi tertentu seperti takut dan marah, dan menunjuk bagaimana emosi diekspresikan ke dalam bahasa, ekspresi wajah, dan *gesture* (bahasa tubuh).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

mempersingkat waktu yang lama untuk mengelompokkan *tweet* tersebut ke dalam masing-masing jenis emosi. Untuk itu dibutuhkan aplikasi agar dapat mempersingkat waktu dalam mengetahui jenis emosi berdasarkan *tweet* tersebut.

Ada banyak metode untuk menyelesaikan persoalan deteksi emosi dari teks. Terdapat 3 pendekatan utama dalam deteksi emosi dari teks yaitu *keyword-spotting*, *statistical*, dan *rule-based* (Bata, Suyoto & Pranowo, 2015). Pendekatan *keyword-spotting* bekerja dengan cara menemukan kata kunci yang berhubungan dengan suatu jenis emosi. Pendekatan *statistical* menggunakan sejumlah besar data (*corpus*) untuk melatih model deteksi emosi. Data (*corpus*) juga digunakan untuk membuat sejumlah aturan dalam pendekatan *ruled-based*.

Pada penelitian-penelitian sebelumnya yang di kembangkan khususnya pada klasifikasi teks untuk mendeteksi emosi yang pernah dilakukan oleh (Sujana,2013) yaitu memanfaatkan big data untuk mendeteksi emosi. Data yang diperlukan sebanyak 5 juta untuk *dataset tweet* dan 7 kategori emosi (sukacita, kesedihan, kemarahan, cinta, rasa takut, rasa syukur, dan kejutan). Metode yang digunakan adalah *Naive Bayes* dan menggunakan *hashtag* untuk mendeteksi emosi. Hasil yang di peroleh menunjukkan bahwa penggunaan *hashtag* dapat diidentifikasi menjadi sebuah emosi yang secara otomatis lebih akurat dibanding metode pendeteksi emosi dengan teks.

Selanjutnya pada penelitian yang dilakukan oleh (Muljono dkk, 2016) mengenai evaluasi dari metode klasifikasi untuk deteksi emosi teks bahasa Indonesia. Penelitian ini mengevaluasi kinerja 4 metode klasifikasi yang berbeda: *Naive Bayes (NB)*, *J48*, *K-Nearest Neighbor (KNN)* dan *Support Vector Machine-Sequential Optimization Minimal (SVM-SMO)*. Penelitian ini menggunakan *corpus* Indonesia, yang berisi 1000 kalimat yang terdiri dari 6 jenis emosi: marah, jijik, takut, gembira, sedih, dan kejutan. Dan hasil yang diperoleh adalah bahwa *SVM-SMO* memberikan kinerja terbaik. Hasilnya menunjukkan bahwa keakuratan *NB*, *J48*, *KNN* dan *SVM-SMO* adalah 80,2%, 80,8%,68,1%, dan 85,5%. Akurasi tertinggi adalah *SVM-SMO* 85,5%.

Pada penelitian sebelumnya yang dilakukan oleh (Aprianti Puji Sujana, 2013) menggunakan 7 kategori emosi (sukacita, kesedihan, kemarahan, cinta, rasa

takut, rasa syukur, dan kejutan), dan pada penelitian (Muljono dkk, 2016) menggunakan 6 kategori emosi (marah, jijik, takut, gembira, sedih, dan kejutan). Pada penelitian ini menggunakan 5 ketegori emosi menurut Power (1997) yang dikutip oleh Destuardi (2009) yaitu, (takut, marah, jijik, sedih, dan senang). Alasan menggunakan 5 jenis emosi karena sejumlah peneliti telah disepakati terdapat 5 jenis dasar emosi manusia (Power, 1997). Penelitian ini akan dilakukan pengklasifikasian emosi pada *tweet* Berbahasa Indonesia menggunakan metode *Support Vector Machine* (SVM).

Support Vector Machine (SVM) merupakan salah satu metode yang banyak digunakan dalam berbagai bidang seperti industri, matematika, statistika, dan teknik informatika. Menurut Widyanto dan Wijayanti (2011) mengatakan bahwa SVM adalah metode klasifikasi paling populer ketiga setelah C4.5 dan K-Means dalam *Internasional Conference on Data Mining* (ICDM 2006). Menurut Pang Ning Tan dkk (2006) yang dikutip oleh Saraswati (2011), teknik klasifikasi SVM telah menunjukkan hasil yang menjanjikan dalam berbagai aplikasi seperti pengenalan digit tulisan tangan sampai kategorisasi teks. Selain itu, SVM juga bekerja sangat baik pada data yang memiliki banyak dimensi. Pemilihan metode SVM menurut (Joachims, 1998) dikutip oleh (Yusra, 2014) menyatakan bahwa SVM adalah metode klasifikasi yang paling tepat untuk melakukan pengklasifikasian teks pada *tweet* dengan tingkat akurasi relatif lebih baik dibanding metode lainnya berdasarkan penelitian yang pernah ia lakukan.

1.2 Rumusan Masalah

Adapun rumusan masalah dari penelitian ini adalah bagaimana mengklasifikasikan *tweet* yang mengandung emosi berbahasa Indonesia dengan menggunakan metode SVM.

1.3 Batasan Masalah

Agar tidak terjadi kesalahan persepsi dalam laporan tugas akhir ini, maka berikut akan di jelaskan beberapa hal yang menjadi batasan masalah pada laporan ini :

1. *Tweet* yang diambil dari Twitter adalah *tweet* yang berbahasa Indonesia.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Klasifikasi jenis emosi yang digunakan adalah takut, marah, jijik, sedih, dan senang.
3. *Dataset* yang digunakan sebanyak 2000 *tweet*.
4. *Tweet* yang dipilih adalah teks dari *tweet* itu sendiri.
5. Kernel yang digunakan yaitu kernel RBF.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai pada penelitian ini adalah menerapkan metode klasifikasi SVM pada *tweet* untuk mengklasifikasikan *tweet* berbahasa Indonesia ke dalam lima kelas emosi yaitu marah, senang, sedih, takut dan jijik.

1.5 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini meliputi:

Tugas akhir memiliki 6 bab, dengan sistematika penulisan laporan sebagai berikut:

BAB I PENDAHULUAN

Bab I ini merupakan bagian yang akan menjelaskan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, dan sistematika penulisan laporan tugas akhir.

BAB II LANDASAN TEORI

Pada bab ini dibahas mengenai pustaka atau literatur yang digunakan dalam pengerjaan skripsi. Teori-teori yang terdapat pada bab ini mencakup tentang text mining, emosi, serta metode *support vector machine*.

BAB III METODOLOGI PENELITIAN

Bab ini berisi tahapan-tahapan yang dilakukan untuk menyelesaikan permasalahan pada tugas akhir seperti tahapan identifikasi masalah, studi literatur, analisa penyelesaian masalah, pengumpulan data, pelabelan manual, *preprocessing*, pembobotan dan seleksi fitur, proses SVM, dan melakukan analisa hasil klasifikasi.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB IV

ANALISA DAN PERANCANGAN

Bab ini berisi tentang analisis dari penelitian yang dilakukan dalam tugas akhir dan menjelaskan tentang perancangan aplikasi.

BAB V

IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi tentang penerapan aplikasi klasifikasi emosi pada *tweet* berbahasa Indonesia seperti tahapan implementasi. Tahapan ini juga menjelaskan pengujian aplikasi dan melakukan perbandingan dengan penelitian terdahulu.

BAB VI

PENUTUP

Bab ini berisi kesimpulan dan saran mengenai hasil analisis, perancangan, hasil implementasi dan hasil pengujian yang telah dilakukan terhadap pengembangan aplikasi klasifikasi emosi pada *tweet*.