

BAB I

PENDAHULUAN

1.1 Latar Belakang

KPU telah menetapkan dua pasangan calon presiden (capres) dan calon wakil presiden (cawapres) Indonesia pada tanggal 20 september tahun 2018 dan pemilu calon presiden beserta calon wakil presiden yang akan diselenggarakan pada tanggal 17 April 2019. Kampanye dilakukan untuk menarik perhatian calon pemilih yang berwenang untuk memilih. Kampanye dilakukan secara daring maupun non daring. Kampanye daring lebih mudah dilakukan karena sangat banyak masyarakat Indonesia yang mempunyai serta menggunakan beberapa akun media sosial, salah satu diantaranya adalah media sosial twitter. Dalam hal ini twitter memiliki peran penting yang dapat digunakan masyarakat maupun partai koalisi untuk berkampanye maupun mencari informasi sentimen tweet dari banyak pihak pendukung maupun pihak lawan.

Untuk memenuhi kebutuhan informasi masyarakat, calon presiden dan calon wakil presiden beserta partai koalisinya akan dilakukan penelitian analisis sentimen data twitter mengenai pendapat masyarakat terhadap capres dan cawapres. Analisa sentimen merupakan bagian dari *opinion mining* [1]. Bidang ini melakukan studi mengenai opini orang-orang, sentimen, evaluasi, tingkah laku dan emosi terhadap suatu entitas seperti produk, layanan, organisasi, individu, permasalahan, topik, acara dan atribut-atributnya[2].

Pada proses klasifikasi sentimen data twitter ada beberapa kasus yang harus di perhatikan yaitu tata bahasa yang digunakan dalam tweet pada twitter dan sentimen sulit untuk diklasifikasikan karena kebebasan pembahasan yang berbeda pada setiap tweet seperti sosial, politik, ekonomi, teknologi, budaya, sains dll.

Penjabaran tersebut akan di gunakan dalam sistem yang akan di buat untuk mengklasifikasi sentimen data twitter dengan melakukan penelitian skripsi berjudul "Analisis Sentimen Data Twitter Terhadap Pemilu 2019 Menggunakan Algoritma Support Vector Machine".

1.2 Rumusan Masalah

Berdasarkan penjabaran latar belakang, terdapat rumusan masalah sebagai berikut:

1. Bagaimana metode *support vector machine* dapat digunakan untuk mengklasifikasi sentimen data twitter?
2. Berapakah akurasi yang dihasilkan dengan menggunakan metode *support vector machine*?

1.3 Batasan Masalah

Berdasarkan penjabaran rumusan masalah, terdapat batasan-batasan dari penelitian yang akan dilakukan antara lain:

1. Data yang digunakan adalah crawling data tweet dari Api twitter mengenai pembahasan pemilu presiden Indonesia tahun 2019.

2. Data yang di ambil merupakan teks berbahasa Indonesia dan tidak menyertai gambar.
3. Data yang diklasifikasi sejumlah 2000 data dan telah di training.
4. Proses *cleansing* dan pemberian label dilakukan secara manual.
5. Metode yang akan di gunakan adalah *support vector machine*.
6. Klasifikasi data dikategorikan menjadi 2 yaitu positif dan negative.
7. Daftar nama calon presiden ada dua yaitu Joko Widodo dan Prabowo Subianto.
8. Bahasa pemrograman yang digunakan untuk membangun system merupakan Python 3.5.
9. Pengujian metode dilakukan dengan menggunakan *cross validation* pada RapidMiner.

1.4 Maksud dan Tujuan Penelitian

1.4.1 Maksud Penelitian

Adapun maksud penelitian ini adalah sebagai berikut:

1. Membuat sistem yang dapat menganalisis data training sentiment twitter yang berhubungan dengan calon presiden Jokowi dan Prabowo.
2. Dapat membantu pasangan calon presiden atau partai koalisi dan masyarakat agar dapat dengan mudah mendapat informasi mengenai sentiment masyarakat terhadap pasangan calon presiden tahun 2019.
3. menyimpulkan hasil klasifikasi dengan menggunakan metode *support vector machine*.

1.4.2 Tujuan Penelitian

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Menganalisis sentiment pada twitter menggunakan algoritma *support vector machine*.
2. Dapat mengumpulkan informasi tweet sentiment masyarakat dengan jumlah yang besar, pada penelitian ini parameter besar adalah lebih dari 1000 tweet sentimen.
3. Dapat menjadi rujukan peneliti selanjutnya yang berhubungan dengan klasifikasi sentiment terutama menggunakan metode *support vector machine*.

1.5 Metodologi Penelitian

Adapun metode-metode penelitian yang akan di gunakan dalam penelitian ini adalah sebagai berikut:

1.5.1 Metode Pengumpulan Data

Metode ini diperlukan untuk mendapatkan informasi yang sesuai dengan topik penelitian. Metode pengumpulan data yang akan dilakukan adalah metode deskriptis dan studi pustaka. Adapun metode pengumpulan data yang di perlukan adalah sebagai berikut:

1. Metode deskriptis

Pada metode deskriptis ini penulis mengumpulkan data dengan cara crawling API twitter mengenai jokowi dan prabowo. Hasil dari pengunduhan sentiment tweet pada twitter merupakan file JSON.

2. Studi pustaka

Penulis melakukan usaha menghimpun informasi dengan cara mengumpulkan referensi penelitian guna menunjang tujuan penelitian. Studi pustaka yang dikumpulkan antara lain buku, jurnal-jurnal, skripsi maupun situs internet mengenai metode *support vector machine*.

1.5.2 Tahapan Penelitian

Tahapan-tahapan penelitian yang dilakukan adalah sebagai berikut:

1. Tahap Analisis

Tahap Analisis digunakan untuk menganalisis dan mempelajari data maupun informasi yang telah didapatkan pada proses pengumpulan data. Kemudian data maupun informasi yang telah dipelajari dan dianalisis digunakan untuk mengembangkan metode yang dapat mengklasifikasi data twitter. Data yang telah didapatkan akan melakukan serangkaian *preprocessing text* dan menganalisis cara kerja dari algoritma SVM. Selanjutnya implementasi metode *support vector machine* dalam bentuk program komputer untuk mengetahui kemampuan metode yang di

usulkan. Setelah proses analisis berhasil, maka akan data akan di uji menggunakan metode *cross validation* di rapidminer.

2. Tahap Perancangan

Metode perancangan yang digunakan untuk membangun sistem pada penelitian ini adalah pembuatan *interface* menggunakan python 3.5, adapun rancangan *interface* yang di akan dibuat yaitu halaman utama yang akan ditampilkan pada halaman awal untuk memilih proses yang akan di jalankan, halaman proses analisis yang akan menampilkan kandidat calon presiden dan untuk memilih halama hasil analisis, dan halaman hasil analisis beserta *piechart* yang akan menampilkan hasil klasifikasi menggunakan metode *support vector machine*.

3. Tahap Implementasi

Pada penelitian ini metode implementasi yang diterapkan antara lain implementasi pengumpulan data, preprocessing data, implmrntasi TFIDF, dan implementasi sistem dengan menggunakan bahasa pemrograman python 3.5 pada *interface* yang akan dibangun untuk mengklasifikasi data menggunakan metode SVM.

4. Tahap Pengembangan

Tahapan metode pengembangan pada penelitian ini adalah sebagai berikut:

1. Crawling API twitter untuk pengambilan data sentiment yang berhubungan dengan jokowi dan prabowo.

2. Melakukan hitungan manual terhadap sample data sentimen yang diambil secara acak.
3. Membuat *user interface* dan kemudian membuat koneksi antara data sentiment twitter atau data trining dengan *interface* program..

5. Tahap Pengujian

Pada tahap ini metode testing diperlukan untuk menguji sistem yang telah dibangun. Dimana pada proses pengujian sistem dapat ditentukan kelayakannya, sehingga hasil testing dapat memastikan bahwa program sesuai dengan kebutuhan fungsionalnya. Metode yang digunakan adalah *cross validation* di rapidminer.

1.6 Sistematika Penulisan

Sistematika Penulisan disusun kedalam beberapa bab sebagai berikut:

BAB I PENDAHULUAN

Pembahasan pada bab ini meliputi latar belakang, rumusan masalah, batasan masalah, maksud dan tujuan, metode penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini memuat tentang teori-teori dan tinjauan pustaka yang digunakan sebagai acuan dasar penelitian.

BAB III ANALISIS DAN PERANCANGAN

Pembahasan dalam bab ini adalah menganalisis kebutuhan data dan perancangan program yang akan digunakan untuk menguji metode yang di usulkan.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pembahasan pada bab ini adalah rincian metode yang diusulkan, implementasi program atau sistem yang akan dibangun serta pengujian metode yang diusulkan.

BAB V PENUTUP

Pembahasan pada penutup yaitu kesimpulan serta saran-saran mengenai penelitian selanjutnya yang dapat dilakukan berkaitan dengan pengolahan data twitter dan metode *support vector machine*.

DAFTAR PUSTAKA

Daftar pustaka memuat tentang referensi yang penulis gunakan sebagai bahan penelitian.