

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Sosial media adalah alat komunikasi atau wadah yang digunakan orang-orang untuk mengemukakan pendapat atau opini mereka untuk berbagai topik. Pengguna sosial media di Indonesia sangat besar, sehingga mendorong munculnya data tekstual yang tidak terbatas. Salah satu pemanfaatan data ini adalah mengetahui sentimen publik terhadap *marketplace*.

Twitter adalah situs jejaring sosial mikroblogging yang digemari oleh masyarakat Indonesia. Survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia menyebutkan bahwa Twitter merupakan media sosial yang menempati peringkat 4 dalam hal kunjungan oleh masyarakat Indonesia. Kelebihan media sosial Twitter adalah dibatasinya jumlah karakter menjadi maksimal 280 karakter, sehingga di twitter tidak ada fitur "*read more*" karena karakter sudah ditampilkan semuanya di layar. *Tweet* yang diposting oleh pengguna tersebut dapat di analisa dan diolah menjadi informasi yang bermanfaat dengan teknik analisa sentimen.

*Text Mining* adalah sebuah teknik untuk menggali informasi dari sebuah teks. Informasi yang diambil melalui teknik *text mining*. *Text mining* merupakan bagian atau variasi dari *data mining*. *Text mining* berusaha menemukan pola dari sekumpulan data. *Text mining* juga dapat diartikan sebagai sebuah proses untuk menemukan suatu informasi atau tren baru yang sebelumnya tidak terungkap dengan memproses dan menganalisis data dalam jumlah besar [1].

*Sentiment analysis* atau *opinion mining* adalah cabang ilmu *text mining* yang mempelajari sentimen yang ada pada suatu teks opini. Prinsip dasar dari analisis sentiment adalah mengklasifikasikan sebuah teks apakah teks tersebut bernilai positif, negatif atau netral. *Sentiment analysis* atau *opinion mining* mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistic dan text mining yang bertujuan menganalisa pendapat, sentiment, evaluasi, sikap, penilaian, dan emosi seseorang [2].

*Marketplace* adalah suatu tempat dimana banyak penjual berkumpul untuk menjual berbagai macam barang produksi yang difasilitasi dengan jaringan internet. Jika dianalogikan, *marketplace* adalah sebuah pasar yang memfasilitasi banyak pedagang dan pembeli untuk melakukan transaksi penjualan. Marketplace merupakan sebuah tren bisnis *e-Commerce* dimana mengajak beberapa pelaku usaha untuk berjualan dalam sebuah *website*, sehingga banyak pembeli bisa memilih barang yang mereka cari dari berbagai macam toko melalui browser atau juga aplikasi mobile. *E-marketplace* adalah wadah komunitas bisnis interaktif secara elektronik yang menyediakan pasar dimana perusahaan dapat ambil andil dalam B2B *e-commerce* dan atau kegiatan *e-business* lain. *E-marketplace* dapat dikatakan sebagai gelombang kedua pada *e-commerce* dan memperluas kombinasi dari bisnis konsumen (B2B,C2B dan C2C) ke dalam B2B. Inti penawaran dari *Marketplace* adalah mempertemukan pembeli dan penjual sesuai dengan kebutuhan dan menawarkan efisiensi dalam bertransaksi [3].

Untuk melakukan proses jual beli saat ini tidak hanya menggunakan cara konvensional seperti pergi ke pasar tradisional, akan tetapi dapat dilakukan secara

elektronik oleh konsumen, dari perusahaan ke perusahaan menggunakan Komputer sebagai media perantara transaksi tersebut. Bukalapak merupakan salah satu marketplace yang sangat populer di Indonesia. Bukalapak memberikan penggunaannya untuk dapat melakukan transaksi secara cepat dan aman. Respon yang diberikan oleh pengguna bukalapak dapat berupa respon positif maupun negatif. Oleh karena itu diperlukan suatu metode yang dapat digunakan untuk mengetahui opini pengguna bukalapak pada sosial media twitter.

Untuk menyelesaikan masalah tersebut maka akan dilakukan penelitian mengenai *text mining* pada sosial media terhadap bukalapak. Dengan penelitian ini dapat menggali opini atau pendapat pengguna bukalapak yang di posting di sosial media apakah itu bersifat positif, negatif maupun netral.

Untuk mewujudkan hal tersebut penulis membuat penelitian dengan judul "Penerapan Metode *Support Vector Machine* Untuk Analisis Sentimen Pengguna Twitter" untuk melakukan klasifikasi opini menggunakan metode support vector machine ini mempertimbangkan pada penelitian-penelitian sebelumnya, menyatakan metode *Support Vector Machine* memiliki akurasi lebih besar dari metode *Naïve Bayes Classifier* [4].

## 1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan, maka dalam penelitian ini permasalahan yang akan dirumuskan diantaranya adalah:

1. Bagaimana Algoritma *Support Vector Machine* dapat digunakan untuk mengklasifikasikan analisis sentimen pengguna Bukalapak berdasarkan situs twitter?

2. Berapa besar tingkat akurasi pemrediksian sentimen terhadap akun twitter dengan menggunakan algoritma *Support Vector Machine* ?

### 1.3 Batasan Masalah

Berdasarkan pembahasan diatas, maka lingkup pembahasan dan perancangan yang akan dibahas hanya dibatasi hal-hal sebagai berikut :

1. Penelitian hanya dilakukan terhadap pengguna media sosial twitter dengan kata kunci Bukalapak..
2. Penelitian ini hanya melakukan analisis sentiment terhadap tweet berbahasa Indonesia.
3. Penelitian ini hanya menggunakan data berupa text saja, tidak mengandung gambar, video, gif, maupun emoticon.
4. Data diunduh dari twitter dari 10-05-2019 sampai 30-07-2019 sebagai data training.
5. Analisa sentimen dengan metode SVM menggunakan tools yang sudah tersedia dan mengklasifikasikan ke dalam tiga kelas sentimen, positif, negatif, dan netral. Pekerjaan yang dilakukan mencakup mengunduh tweet, tweet preprocessing, klasifikasi tweet, pengujian data, serta pemanfaatan tools atau library yang tersedia untuk analisis sentiment maupun tweet preprocessing.
6. Bahasa pemrograman yang digunakan adalah Python.

### 1.4 Maksud dan Tujuan Penelitian

Tujuan diadakannya penelitian ini adalah :

1. Mengimplementasikan algoritma *Support Vector Machine* dalam sistem analisis sentiment untuk mengklasifikasikan opini pengguna twitter.
2. Mencari nilai akurasi Algoritma Support Vector Machine dalam mengklasifikasi sentimen terhadap pelayanan marketplace Bukalapak.

### 1.5 Manfaat Penelitian

#### 1. Bagi Pengguna

Pengguna dapat mengetahui bagaimana performa bukalapak berdasarkan opini yang telah diberikan oleh pengguna pada sosial media twitter.

#### 2. Bagi Penulis

- a. Menambah serta mengasah ilmu yang telah dipelajari di Universitas Amikom Yogyakarta.
- b. Sebagai Syarat Kelulusan program studi Sarjana S1 jurusan Informatika Universitas Amikom Yogyakarta.
- c. Pembuatan Karya Ilmiah ini sebagai bukti turut berperan dalam bidang pengembangan Teknologi.

#### 3. Bagi Bukalapak

Pengklasifikasian data tweet tentang Bukalapak dapat memudahkan pihak perusahaan dalam mengetahui persepsi pengguna jasa dalam bentuk opini negatif dan opini positif, sehingga dapat dijadikan sebagai acuan dalam upaya menjaga kualitas dan memperbaiki kekurangan serta evaluasi ke arah yang lebih baik.

## 1.6 Metode Penelitian

Metodologi yang digunakan dalam pengerjaan tugas akhir ini adalah metode *experimental* dan dilaksanakan dalam tahapan-tahapan berikut:

### 1.6.1 Metode Studi Literatur

Penulis mempelajari berbagai macam literatur, khususnya mengenai algoritma *Support Vector Machine*. Penulis juga mengumpulkan berbagai macam referensi seperti jurnal penelitian, tesis, buku-buku teori, dan sumber-sumber lain tidak terkecuali sumber yang berasal dari internet.

### 1.6.2 Metode Analisis

Dalam tahap ini, penulis melakukan analisa terhadap tahap demi tahap dari proses *text mining* hingga mencapai hasil akhir.

### 1.6.3 Perancangan

Pada tahap ini dilakukan perancangan metode kombinasi antara *featureselection* menggunakan POS Filtering dan SVM yang akan diimplementasikan untuk menganalisa sentimen. Data-data yang diambil dalam penelitian ini diambil dari Twitter menggunakan *twitterscraper*, selanjutnya data tersebut disimpan dalam file dokumen. Perancangan sistem meliputi training dan testing. Training sendiri terdapat beberapa tahap meliputi melabeli *tweet* secara manual, *preprocessing*, dan proses *training* sendiri. Pertama-tama, secara manual, data training akan dikelompokkan menjadi dalam 3 kelompok, positif, negatif, dan netral. Lalu dilanjutkan ke tahap *preprocessing*.

Setelah melakukan tahap *preprocessing* maka data tweet yang sudah diberi label serta sintaks masing-masing kata dihitung bobotnya. Kemudian, bobot

tersebut akan disimpan dalam berkas. Selanjutnya akan dilakukan proses pelatihan dengan memanfaatkan *library Scikit-Learn*. Hasil dari pelatihan ini adalah SVM model yang disimpan dalam berkas.

Pada proses testing akan menggunakan metode SVM yang memanfaatkan *library Scikit-Learn*. Dokumen yang berisikan *tweet* akan dianalisa sentimennya melalui SVM model yang telah dibuat. Selanjutnya, dilakukan *preprocessing* terhadap data *tweet* yang didapat. Lalu, data *tweet* akan diidentifikasi bobotnya dengan melakukan pembobotan *Cosine Similarity* terhadap *tweet* tersebut. Langkah terakhir adalah ditentukan labelnya melalui klasifikasi yang dilakukan oleh model SVM yang ada dalam *library Scikit-Learn*.

#### **1.6.4 Implementasi**

Pada tahap ini akan dilakukan pembuatan kode program sampai proses klasifikasi sentimen dokumen.

#### **1.6.5 Pengujian**

Tahap pengujian yaitu tahap perhitungan akurasi dari implementasi yang telah dilakukan sebelumnya. Metode yang digunakan untuk menguji yaitu *K-Fold Cross Validation*. Pengujian yang dilakukan adalah perbandingan akurasi klasifikasi yang dihasilkan *Support Vector Machine*.

## 1.7 Sistematika Penulisan

Sistematika penulisan laporan penelitian ini adalah sebagai berikut :

### **BAB I : PENDAHULUAN**

Menjelaskan tentang latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, metode penelitian dan sistematika penulisan laporan.

### **BAB II : LANDASAN TEORI**

Menjelaskan teori-teori yang mendasari pembahasan dalam penyusunan skripsi.

### **BAB III : PERANCANGAN**

Pada bab ini akan membahas tentang desain dan perancangan sistem.

### **BAB IV : IMPLEMENTASI DAN PENGUJIAN**

Bab ini membahas hasil tentang implementasi dan pengujian sistem

### **BAB V : PENUTUP**

Bab ini berisi kesimpulan dan saran dari apa saja yang telah dijelaskan pada bab sebelumnya.