

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini jasa antar barang menjadi sarana prasarana yang memudahkan masyarakat dalam beraktifitas. Salah satu startup dibidang jasa pengantar barang ini adalah JNE. JNE merupakan suatu perusahaan dibidang jasa yang bertujuan untuk mempermudah masyarakat dalam melakukan pengantaran barang kemanapun yang ingin diinginkan oleh pelanggan. JNE juga menyediakan aplikasi berbasis mobile yang bertujuan untuk mempermudah masyarakat untuk menkira – kira ongkos yang dikeluarkan untuk jasa pengantaran barang ini dan bisa juga untuk melacak barang sudah sampai tujuan atau belum dengan memasukkan nomor resinya. Peran konsumen JNE dalam memberikan opini berupa kritik dan saran dapat berguna sebagai feedback untuk Perusahaan JNE agar bisa lebih baik dalam memberikan pelayanannya.

Sabagian besar konsumen mengambil saran, ulasan ataupun pendapat dari opini orang lain sebelum mereka memutuskan untuk membeli barang atau menggunakan jasa tertentu. Media sosial adalah tempat dimana seseorang dapat mengemukakan sentimen, pendapat, ulasan, kritik tanpa ada yang membatasi opini mereka. Dampak dari media sosial ini sangat berpengaruh baik untuk konsumen maupun produsen tanpa mereka sadari.

Dengan demikian, dari ulasan – ulasan tersebut dapat memberikan nilai potensial yang dapat diimplementasikan untuk analisis sentimen yang bertujuan untuk meningkatkan strategi dalam pemasaran, peningkatan produk, pelayanan untuk menjadi lebih baik. Salah satu contoh media sosial tempat bertukar opini adalah Twitter. Dari twitter dapat dilakukan Opinion Mining atau bisa disebut Analisis Sentimen. Untuk melakukan Twitter Mining peneliliti akan mengambil data yang digunakan dengan memanfaatkan API Twitter.

Analisis sentimen merupakan suatu proses pengolahan data berupa teks yang bertujuan untuk mengklasifikasikan sentimen positif. Dalam proses pengolahan data sentimen akan melewati beberapa tahapan yaitu Pengumpulan Data, Cleaning

Data, Tokenization, Stopword Removal, Stemming, Feature List, Pembobotan kemudian setelah itu diklasifikasikan menggunakan algoritma Support Vector Machine. SVM merupakan salah satu metode klasifikasi dengan menggunakan metode machine learning (supervised learning) yang memprediksi kelas berdasarkan pola dari hasil proses training yang diciptakan oleh Vladimir Vapnik. Klasifikasi dilakukan dengan garis pembatas (hyperplane) yang memisahkan antara kelas opini positif dan opini negatif [1]. Dalam penelitian yang berjudul "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 di Twitter" yang dilakukan oleh Ghulam Asrof Buntoro [2] dapat disimpulkan bahwa metode ini adalah metode yang paling baik dibandingkan dengan metode yang lainnya. Dengan menggunakan metode Support Vector Machine, tingkat keakuratan akan mencapai 95%. Oleh karena itu maka dalam penelitian ini akan menggunakan metode Support Vector Machine untuk mengukur nilai akurasi sentimen yang diperoleh dari hasil crawling data twitter JNE.

1.2 Rumusan Masalah

Perumusan masalah dalam skripsi ini adalah:

1. Bagaimana Algoritma Support Vector Machine menemukan sentimen yang diberikan oleh pengguna jasa antar barang JNE dari komentar twitter
2. Bagaimana nilai akurasi algoritma Support Vector Machine pada klasifikasi sentimen terhadap pelayanan jasa antar barang JNE

1.3 Batasan Masalah

Batasan masalah dalam pembuatan Skripsi ini adalah:

1. Penelitian hanya dilakukan terhadap data twitter dengan hastag yang berhubungan dengan JNE seperti '*jne*', '*JNECare*', '*jneindonesia*', '*YESJNE*'.
2. Bahasa yang diambil dari Twitter adalah Bahasa Indonesia.
3. Analisis sentimen dalam penelitian ini menggunakan data tweet dari Twitter yang diunduh selama tiga kali.

4. Analisis sentimen dengan metode SVM menggunakan tools yang sudah tersedia di library python.
5. Dari sentimen diklasifikasikan kedalam 2 kelas yaitu positif dan negatif.
6. Bahasa pemrograman yang digunakan adalah Python.
7. Training data dilakukan secara terpisah dengan test predik suatu data.
8. Analisis sentimen untuk data pengujian data testing dalam penelitian ini menggunakan data tweet dari Twitter yang diunduh secara *realtime*.
9. Jumlah data training 800 data

1.4 Maksud dan Tujuan Penelitian

Tujuan dari pembuatan Skripsi ini adalah:

1. Menerapkan metode SVM pada analisis sentimen terhadap data twitter JNE.
2. Menemukan sentimen yang berhubungan dengan JNE dari komentar twitter menggunakan algoritma Support Vector Machine.
3. Mencari nilai akurasi Algoritma Support Vector Machine dalam mengklasifikasi sentimen terhadap tingkat kepuasan Pelayanan JNE.

1.5 Manfaat Penulisan

- a. Bagi Perusahaan JNE
Mendapatkan presentase sentimen publik dari data twitter JNE terhadap pelayanan dari JNE guna meningkatkan pelayanan dan mutu dari JNE.
- b. Bagi Ilmu Pengetahuan
 1. Mengetahui bagaimana implementasi opinion mining dalam pemrograman python.
 2. Mengetahui cara memperoleh, mengolah data untuk sentimen analisis.

1.6 Metode Penelitian

Dalam penelitian ini akan menjelaskan tahapan – tahapan sentimen analisis sebagai berikut:

1.6.1 Metode Pengumpulan Data

1.6.1.1 Studi Pustaka

Mencari literatur atau jurnal yang bersangkutan dengan Analisa Sentimen, Opinion Mining, Algoritma Support Vector Machine, K-Fold Cross Validaton.

1.6.1.2 Survei

Mangamati kebutuhan koresponden dengan melakukan penyebaran suatu pernyataan yang nantinya menjadi dasar dalam pemberian suatu label pada data Twitter yang akan dilatih.

1.6.1.3 Eksperimen

Menguji kode program agar mudah untuk dikembangkan dan mencari metode penulisan program yang tepat. Membuat *user interface* yang mudah untuk dipahami pengguna serta dapat memberikan informasi kepada pengguna.

1.6.2 Metode Analisis

Dalam metode analisis ini terdapat beberapa tahapan seperti dibawah ini:

1. Collecting Data

Data yang dikumpulkan diperoleh dari crawling data twitter menggunakan Twitter API.

2. Data Preprocessing

Tahap *preprocessing* adalah tahap awal dalam analisis sentimen. Berikut tahap – tahapnya :

- a. Cleaning Data merupakan tahapan penghapusan kata – kata yang tidak dipakai seperti hastag dan username.
- b. Case Folding atau Lower Case digunakan untuk membuat keseluruhan data tweet menjadi lowercase (huruf kecil)
- c. Tokenization merupakan tahap untuk memisahkan kalimat menjadi kata dan menghilangkan data yang tidak bermakna.

- d. Stemming merupakan tahapan untuk menghilangkan imbuhan awal atau akhir pada kata tersebut seperti berlari menjadi lari.
3. Feature Weighting (Fitur Pembobotan)
 - a. *Term Frequency (TF)* merupakan frekuensi kemunculan kata pada kalimat.
 - b. *Document Frequency (IDF)* kemunculan berapa banyak kata dalam satu kalimat.
4. Klasifikasi dengan feature Cosine Similarity
Feature Cosine Similarity adalah fitur untuk mencari tingkat kesamaan data untuk meningkatkan akurasi data.
5. Klasifikasi Support Vector Machine
Pada tahap ini data diklasifikasikan menggunakan Algoritma Support Vector Machine dengan menentukan suatu kelas berdasarkan pola yang terbentuk dari hasil proses training. Disini data klasifikasi menjadi 2 tahap yaitu dengan tahap training yang kemudian akan diberikan label secara manual dan untuk data pengujian yang diunduh secara realtime.
6. Menghitung akurasi dari pengujian data
Untuk menguji seberapa akurat penelitian ini maka akan dilakukan pengujian dengan 10-fold sebagai parameternya untuk melakukan Cross Validation yang diuji menggunakan tools pada python

1.6.3 Metode Perancangan

1. Perancangan Aplikasi menggunakan flowchart
2. Perancangan Antarmuka Pengguna

1.6.4 Metode Implementasi

1. Implementasi Sistem di shell python dengan menggunakan bahasa python.
2. Manual Program
3. Hasil Implementasi

1.6.5 Metode Testing

K-fold merupakan metode yang digunakan untuk mengetahui rata-rata keberhasilan suatu sistem.

1.7 Sistematika Penulisan

Sistematika dalam penulisan skripsi ini dibagi menjadi lima bab, antara lain sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini akan menjelaskan mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan, manfaat dan sistematika penulisan penelitian.

BAB II LANDASAN TEORI

Pada bab ini menjelaskan mengenai teori yang digunakan sebagai dasar dalam analisis penelitian. Mencakup tinjauan pustaka meliputi Pengertian *Data Mining*, *Sentiment Analysis*, *Support Vector Machine (SVM)*, *Pre-Processing*, *Term Frequency-Inverse Document Frequency (TF-IDF)*, *Cosine Similarity*, *K-Fold Cross Validation* dan *Confusion Matrix*.

BAB III ANALISIS DAN PERANCANGAN

Pada bab ini akan menjelaskan mengenai perancangan dan model cara kerja yang akan dilakukan di dalam studi kasus sentimen analisis tweet JNE.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini akan menjelaskan *Analisis Sentimen* dari *Pre-Processing* data, *Tokenization*, *Steaming*, *TF-IDF*, *SVM* dan *Confusion Matrix*.

BAB V PENUTUP

Bab ini akan menjelaskan kesimpulan yang didapat dari hasil analisis penelitian ini, perancangan dan website serta saran untuk pengembangan yang lebih baik lagi.

