

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini persaingan antar maskapai penerbangan semakin hari semakin banyak. Masing-masing dari maskapai penerbangan berlomba-lomba untuk mendapatkan pelanggan sebanyak-banyaknya. Semakin banyak pengguna tentunya akan menguntungkan bagi pihak maskapai penerbangan. Namun kualitas maskapai penerbangan tidak dapat ditentukan dari jumlah penggunanya saja. Tetapi seberapa luas infrastruktur yang dimiliki oleh maskapai penerbangan, semakin luas jangkauan maka probabilitas pelanggan semakin tinggi.

Dalam industri transportasi layanan tentunya tidak menjamin semua pelanggan puas dengan layanan yang diberikan, pasti ada opini dari pelanggan yang bermacam-macam. Di jaman modern ini pelanggan biasanya meluapkan opini positif maupun negatif di sosial media mereka, karena mereka mempunyai kontrol penuh terhadap sosial media sendiri.

Dengan itu opini-opini yang muncul bisa memberikan nilai positif terhadap pelayanan dari maskapai penerbangan yang ada di Indonesia untuk meningkatkan pelayanan, strategi pemasaran, dan peningkatan produk agar lebih baik lagi. Salah satu media sosial yang digunakan tempat beropini yaitu twitter. Dari twitter dapat dilakukan *opinion mining* atau yang disebut analisis sentimen. Untuk mendapatkan data guna melakukan twitter mining peneliti akan memanfaatkan *API Twitter* untuk mengambil data.

Penulis bermaksud untuk membuat sebuah aplikasi analisis sentimen pengguna twitter terhadap pelayanan maskapai penerbangan berdasarkan opini masyarakat di twitter dengan berbasis sentiment analisis menggunakan algoritma *Support Vector Machine* , dengan tujuan mengetahui presentase sentiment positif dan negatif dari maskapai penerbangan di Indonesia sehingga bisa maskapai penerbangan di Indonesia bisa meningkatkan pelayanan.

1.2 Rumusan Penelitian

Berdasarkan uraian latar belakang permasalahan di atas , maka penulis mengidentifikasi masalah sebagai berikut :

1. Bagaimana melakukan analisis sentimen yang diberikan oleh pengguna maskapai penerbangan dari *tweet* yang ada di twitter menggunakan algoritma *Support Vector Machine*?
2. Bagaimana mengukur nilai akurasi dengan *K-Fold Cross Validation* dan *Confusion Matrix* dari algoritma *Support Vector Machine* pada klasifikasi sentimen terhadap pelayanan maskapai penerbangan yang ada di Indonesia?

1.3 Batasan Penelitian

Dari rumusan masalah diatas maka dibuatlah Batasan masalah agar penulisan skripsi ini tidak menyimpang dari tujuan yang direncanakan:

1. Bahasa yang diambil dari twitter adalah Bahasa Indonesia.

2. Penelitian hanya dilakukan pada *tweet* yang memiliki hastag yang berhubungan dengan maskapai penerbangan seperti '#lionair', 'Lion Air', '#garudaindonesia', 'Garuda Indonesia'.
3. Bahasa pemrograman yang dilakukan adalah Python dengan versi 3.10
4. Analisa sentimen menggunakan *library* yang sudah tersedia di Python dan menggunakan *Support Vector Machine*.
5. Data sentimen diklasifikasikan kedalam 2 kelas yaitu positif dan negatif.
6. Data yang digunakan untuk pengujian dalam penelitian ini menggunakan data *tweet* dari twitter.
7. Jumlah data training 700 data.
8. Perhitungan akurasi dari algoritma *Support Vector Machine* menggunakan *K-Fold Validation (10-fold)* dan *Confusion Matrix*.

1.4 Maksud dan Tujuan Penelitian

1.4.1 Maksud Penelitian

Menerapkan metode *Support Vector Machine* untuk Analisa sentimen terhadap pelayanan maskapai penerbangan.

1.4.2 Tujuan Penelitian

1. Menemukan sentimen yang berhubungan dengan maskapai penerbangan di twitter menggunakan algoritma *Support Vector Machine*.
2. Mencari nilai akurasi algoritma *Support Vector Machine* dalam klasifikasi sentiment terhadap pelayanan maskapai penerbangan dengan menggunakan *K-Fold Validation (10-fold)* dan *Confusion Matrix*.

1.5 Metodologi Penelitian

Metode penelitian yang dilakuakn untuk mendapatkan informasi dan data mengenai objek permasalahan ini adalah :

1.5.1 Metode Pengumpulan Data

Metode pengumpulan data yang di gunakan sebagai berikut:

1. Studi Pustaka

Mencari literatur atau jurnal yang bersangkutan dengan analisis sentimen, *opinion mining*, *algoritma Support Vector Machine*, *K-Fold Cross Validation*, dan *Confusion Matrix*.

2. Metode Wawancara

Perlunya pengetahuan mengenai langkah-langkah yang digunakan dalam analisis sentimen, sehingga dibutuhkan wawancara bersama orang yang ahli atau sudah pernah melakukan analisis sentimen, selain itu dibutuhkan juga informasi mengenai cara kerja dari metode-metode yang biasa digunakan dalam analisis sentimen.

1.5.2 Tahapan Penelitian

1. Analisis

Analisis pada tahapan penelitian ini terdapat beberapa tahapan, yaitu sebagai berikut :

a. *Collecting Data*

Data diperoleh dari crawling data twitter menggunakan API Twitter selama 4 kali dalam rentang waktu 4 minggu.

b. *Preprocessing Data*

Merupakan tahap awal pengolahan data dalam data mining, yang tahapannya terdiri dari :

1. *Data Cleaning*, adalah tahap menghapus kata yang tidak dipakai seperti hastag dan username, dan juga mengurangi noise saat menganalisis sentimen.
2. *Case Folding*, merupakan cara untuk membuat keseluruhan data menjadi lowercase (huruf kecil).
3. *Tokenization* merupakan tahap memisahkan kalimat menjadi kata.
4. *Stopword* merupakan tahap yang berfungsi menghapus kata yang tidak memiliki arti yang deskriptif.

c. *Feature Weighting*

Feature Weighting merupakan proses pembobotan setiap kata pada setiap dokumen, pada penelitian ini menggunakan Teknik *TF-IDF*.

d. *Cosine Similarity*

Merupakan fitur yang digunakan untuk mencari tingkat kesamaan data untuk meningkatkan nilai akurasi.

e. *Klasifikasi Support Vector Machine*

Pada tahap ini data akan dilatih menggunakan Algoritma *Support Vector Machine* dengan menentukan suatu kelas berdasarkan pola yang terbentuk dari hasil proses *training*. Klasifikasi dilakukan dengan memisahkan antara opini positif dan opini negatif.

f. Menghitung akurasi model

Untuk menguji keakuratan model dari *Support Vector Machine* maka dilakukan pengujian menggunakan *confusion matrix* agar menghasilkan performa model berupa nilai akurasi, *precision*, *recall*, dan *f1-score* dengan *10-Fold Cross Validation* sebagai metode untuk membagi data *training* dan data *testing*, dengan hasil akhir yaitu nilai rata-rata seluruh *fold*.

2. Perancangan

Perancangan yang perlu dilakukan adalah perancangan untuk mengimplementasikan algoritma ke dalam aplikasi menggunakan *flowchart*. Penerapan algoritma *Support Vector Machine* untuk suatu sistem diperlukan perancangan sistem dari tahap *pre-processing*, *training*, hingga *testing*, serta alur dari *Support Vector Machine* bekerja. Selain itu untuk memudahkan penggunaan aplikasi perlu dirancang nya antarmuka pengguna.

3. Implementasi

Proses implementasi pembuatan sistem di lakukan menggunakan bahasa pemrograman Python dengan memanfaatkan *library* dari *sklearn*, *numpy*, *pandas*, *sastrawi*, dan *nlTK*.

4. Pengujian

Pengujian dilakukan dengan menggunakan metode *K-fold Cross Validation* yang merupakan teknik untuk melakukan evaluasi kinerja model. Pengujian membagi kedalam data *training* dan data *testing*

menggunakan metode *10-Fold Cross Validation*. Masing-masing *fold* akan dihitung performanya menggunakan metode *confusion matrix* untuk memberikan hasil akhir performa model.

1.6 Sistematika Penulisan

Laporan ini akan disusun secara sistematis dalam bentuk bab dan masing-masing bagian akan diuraikan ke dalam beberapa sub-bab agar yang akan diurutkan sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini akan menjelaskan mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan, manfaat, metode penelitian 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*, *Support Vector Machine (SVM)*, analisis sentimen, *Pre-Processing*, *Term Frequency - Inverse Document Frequency*, *Cosine Similarity*, *K-Fold Cross Validation* dan *Confusion Matrix*.

BAB III METODE PENELITIAN

Pada bab ini akan menjelaskan mengenai perancangan model, cara kerja yang akan dilakukan didalam analisis sentimen twitter maskapai penerbangan.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bagian bab ini akan menjelaskan kode program mulai dari *pre-processing*, hingga mendapatkan model algoritma *Support Vector Machine*, dan *testing*. Serta menerapkan dari model algoritma *Support Vector Machine* ke dalam aplikasi.

BAB V PENUTUP

Pada bab ini berisi kesimpulan yang di dapat dari hasil analisis, perancangan dan implementasi yang telah dibahas pada bab-bab sebelumnya serta saran untuk pengembangan yang lebih baik.

