

## **BAB I PENDAHULUAN**

### **1.1 Latar Belakang**

Dalam konteks globalisasi saat ini, khususnya di Indonesia yang memiliki populasi yang padat, kebutuhan akan layanan transportasi semakin kompleks seiring dengan meningkatnya mobilitas dan aktivitas manusia. Dengan adanya pembangunan tol Trans Jawa, yang merupakan proyek jaringan jalan tol yang menghubungkan berbagai kota di Pulau Jawa, mulai dari pelabuhan Merak hingga Probolinggo di Jawa Timur, tidak hanya pengguna kendaraan pribadi dan truk ekspedisi yang diuntungkan, tetapi juga perusahaan otobus. Dampak positif dari tol Trans Jawa ini bagi pengusaha otobus meliputi waktu tempuh yang semakin cepat dan efisien, serta harga tiket yang menjadi relatif lebih murah dibandingkan dengan transportasi darat lainnya.[1]

Transportasi merupakan sebuah sistem yang terdiri dari sarana dan prasarana yang didukung oleh pengelolaan yang efisien serta tenaga kerja yang berkualitas. Padatnya penduduk dan kompleksitas kebutuhan memicu persaingan yang ketat di antara berbagai moda transportasi, baik darat, laut, maupun udara. Persaingan ini mendorong perusahaan transportasi untuk bersaing dalam memberikan layanan terbaik guna menarik minat penumpang dan mempertahankan eksistensinya. Salah satu moda transportasi yang sangat populer di kalangan masyarakat Indonesia adalah bus [2].

Pada tahun 2022, menurut survei yang dilakukan oleh Balitbang Kemenhub (2022), bus menjadi moda transportasi umum yang paling sering digunakan oleh penduduk, dengan total 14,1 juta penumpang. Jumlah ini lebih tinggi dibandingkan dengan pesawat yang memiliki 8,9 juta penumpang, dan kereta api yang mengangkut 7,6 juta penumpang. Data ini menunjukkan bahwa hingga saat ini, bus masih menjadi moda transportasi umum dengan tingkat popularitas tertinggi dibandingkan moda transportasi lainnya. Seiring dengan meningkatnya jumlah

penumpang bus, panjang jalan tol di Indonesia juga terus bertambah setiap tahunnya.[3]

Twitter merupakan sebuah platform media sosial yang sangat populer di kalangan pengguna internet. Pengguna Twitter dapat mengekspresikan pikiran, pendapat, dan ide-ide mereka dalam bentuk pesan yang singkat secara langsung dan real-time melalui layanan microblogging yang disediakan[4]. Pengguna Twitter bebas menyampaikan pendapat atau ekspresi mereka tentang layanan, fasilitas atau isu politik atau hal-hal yang sedang dibahas.

Dalam penelitian ini, analisis sentimen menggunakan dua teknik klasifikasi yang dikenal sebagai metode *K-Nearest Neighbor* dan *Support Vector Machine*. Pemilihan metode *K-Nearest Neighbor* didasarkan pada sifatnya yang dianggap relatif sederhana namun tetap efektif dalam mengklasifikasikan teks. Metode ini dianggap sederhana karena memiliki kemampuan untuk mengenali pola secara otomatis dari data yang ada dan mengklasifikasikan data berdasarkan pada tetangga terdekat[5]. *Support Vector Machine* dipilih karena mampu menangani data berdimensi tinggi dan mengatasi ketidakseimbangan kelas dengan baik, serta memiliki dukungan implementasi yang tersedia di Python[6].

Oleh karena itu, penelitian ini bertujuan untuk menggali lebih dalam tentang persepsi dan sentimen masyarakat Indonesia terhadap PO Rosalia Indah di platform *Twitter*. Analisis sentimen yang dilakukan akan memberikan wawasan yang berharga tentang preferensi, harapan, dan kebutuhan masyarakat terkait transportasi darat, serta bagaimana mereka mengevaluasi kualitas layanan yang disediakan.

Dengan pemahaman yang mendalam tentang sentimen masyarakat, PO Rosalia Indah dapat mengidentifikasi area-area yang memerlukan perbaikan atau peningkatan dalam layanan mereka. Selain itu, hasil penelitian ini juga diharapkan dapat memberikan informasi yang berharga bagi masyarakat tentang pengalaman nyata dalam menggunakan layanan transportasi, serta membantu perusahaan dalam meningkatkan transparansi dan akuntabilitas dalam menyediakan layanan yang berkualitas.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah yang dapat diselesaikan dalam penelitian ini adalah:

1. Bagaimana sentimen masyarakat Indonesia terhadap PO Rosalia Indah berdasarkan komentar di Twitter?
2. Bagaimana penerapan algoritma KNN dalam menganalisis sentimen pengguna Twitter terhadap PO Rosalia Indah?
3. Bagaimana penerapan algoritma SVM dalam menganalisis sentimen pengguna Twitter terhadap PO Rosalia Indah?
4. Algoritma mana yang menghasilkan klasifikasi sentimen yang lebih akurat antara KNN dan SVM dalam menganalisis komentar terkait PO Rosalia Indah di Twitter?

## 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini meliputi aspek-aspek berikut:

1. Data yang digunakan dalam penelitian ini berasal dari Twitter dengan kata kunci "rosalia indah" dan "Po rosalia indah," dikumpulkan melalui proses crawling.
2. Metode klasifikasi yang diterapkan dalam penelitian ini adalah K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM), yang akan dibandingkan berdasarkan akurasi klasifikasi sentimen.
3. Penelitian ini difokuskan pada analisis sentimen dari pengguna Twitter mengenai PO Rosalia Indah, dengan subjek penelitian berupa pengguna Twitter yang berbahasa Indonesia.
4. Analisis sentimen ini akan mengklasifikasikan komentar pengguna ke dalam tiga kategori: positif, negatif, dan netral.

5. Penelitian ini menggunakan bahasa pemrograman Python dan memanfaatkan Google Collaboratory sebagai platformnya.

#### **1.4 Tujuan Penelitian**

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah:

1. Mengelompokkan komentar-komentar terkait PO Rosalia Indah di Twitter ke dalam sentimen positif, netral, dan negatif dengan menggunakan metode K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM).
2. Algoritma mana yang menghasilkan klasifikasi sentimen yang lebih unggul akurasi antara KNN dan SVM dalam menganalisis komentar terkait PO Rosalia Indah di Twitter?

#### **1.5 Manfaat Penelitian**

Adapun manfaat dari penelitian berikut adalah:

1. Memberikan wawasan kepada PO Rosalia Indah mengenai persepsi masyarakat, khususnya di platform Twitter.
2. Memberikan wawasan mengenai efektivitas algoritma KNN & algoritma SVM dalam analisis sentimen.
3. Memberikan sumber referensi untuk penelitian selanjutnya terkait analisis sentimen.

#### **1.6 Sistematika Penulisan**

Berisi Berikut ini merupakan sistematika penulisan yang ada dalam penelitian ini:

**BAB I PENDAHULUAN:** Berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian serta sistematika penulisan skripsi.

**BAB II TINJAUAN PUSTAKA:** Berisi studi literatur, dasar-dasar teori yang digunakan tersebut akan dijadikan acuan oleh penulis dalam menyelesaikan

permasalahan yang dibahas.

**BAB III METODE PENELITIAN:** Di dalamnya membahas metode yang dilakukan dalam penelitian serta Langkah-langkah yang harus dilakukan untuk mencapai tujuan dalam menyelesaikan penelitian tugas akhir.

**BAB IV HASIL DAN PEMBAHASAN:** Bab ini merupakan tahapan proses yang dilakukan dalam penelitian, mulai dari pengumpulan data hingga hasil yang didapatkan dari penelitian.

**BAB V PENUTUP:** Berisi kesimpulan dan saran yang dapat peneliti rangkum selama proses penelitian yang diharapkan dapat bermanfaat untuk pengembangan dan melanjutkan penelitian.

