

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Teknologi berkembang dengan cepat, begitu pula aplikasi di *smartphone* yang menggunakan sistem operasi seperti Android atau iOS. Pertumbuhan teknologi yang pesat ini telah memberikan dampak signifikan pada berbagai sektor bisnis, termasuk sektor transportasi umum. Kini, terdapat banyak variasi aplikasi yang menawarkan layanan serupa di sektor ini. Aplikasi-aplikasi tersebut dapat dengan mudah diunduh melalui platform seperti *Google Play Store* dan *App Store*. Pertumbuhan teknologi pada aplikasi mobile menghasilkan data berharga yang, jika diolah, dapat memberikan informasi penting bagi perusahaan. Kepuasan dan loyalitas pelanggan adalah salah satu tujuan utama perusahaan. Melalui pengukuran kepuasan pelanggan, perusahaan dapat mengevaluasi servis atau produk saat ini dan menggunakan hasil evaluasi tersebut untuk memberikan rekomendasi yang membantu meningkatkan servis atau produk perusahaan [1].

Kereta Cepat Jakarta-Bandung (KCJB) adalah proyek infrastruktur transportasi yang dikelola oleh PT Kereta Cepat Indonesia China (KCIC). Proyek ini bertujuan untuk menyediakan solusi transportasi yang lebih cepat dan nyaman untuk masyarakat yang ingin berpergian dari Jakarta atau Bandung. Diluncurkan pada tahun 2015, proyek ini mencapai tahap komersial dengan diluncurkannya penjualan tiket melalui situs web dan aplikasi Whoosh pada tanggal 17 Oktober 2023 [2]. Aplikasi resmi oleh PT KCIC ini dibuat untuk mempermudah masyarakat mengakses informasi dan layanan KCJB secara online, tersedia di *Google Play Store* dan *App Store*. Namun, menurut data dari *Google Play Store*, aplikasi KCJB hanya mendapatkan rating 2,5 dan 1,7 di *App Store*. Rating dari kedua platform tersebut tergolong rendah mengingat pentingnya sebagai aplikasi resmi dari proyek infrastruktur besar yang menjadi bagian dari proyek strategis nasional.

Berdasarkan nilai *rating* rendah di *Google Play Store* dan *App Store* terdapat indikasi adanya ketidakpuasan atau masalah yang dihadapi pengguna. Komentar-komentar ini memberikan gambaran yang mewakili keluhan dan kesukaan yang sering muncul, serta mencerminkan pandangan pengguna terhadap pengalaman yang sering mereka alami saat menggunakan aplikasi Whoosh. Penting untuk disadari bahwa banyaknya ulasan dari kedua *platform* sulit untuk diolah dan

dianalisis secara manual dalam menentukan sentimen masyarakat terhadap aplikasi. Meskipun nilai rating dapat menjadi indikator yang membantu, namun rating saja tidak cukup untuk memberikan pemahaman yang mendalam terhadap keluhan pengguna pada aplikasi.

Selain itu, variasi dalam bahasa dan gaya penulisan pengguna juga menambah kompleksitas dalam analisis. Oleh karena itu, diperlukan sebuah mesin yang dapat melakukan analisis sentimen secara otomatis dan mengklasifikasikan pendapat pengguna menjadi negatif dan positif berdasarkan ulasan yang diberikan. Salah satu teknik untuk menentukan pandangan/opini seseorang pada suatu hal yaitu teknik analisis sentimen adalah metode dalam pengolahan bahasa alami yang berfungsi untuk mengenali opini dan perasaan seseorang terhadap suatu isu atau topik spesifik. Proses ini menganalisis data untuk mengetahui kecenderungan opini, apakah lebih condong ke arah positif atau negatif, berdasarkan reaksi dan tanggapan terhadap objek tertentu [3]. Meskipun tantangan dalam memproses data yang besar dan beragam, teknik analisis sentimen mampu memberikan pemahaman yang berharga terhadap suatu hal untuk memahami sentimen pengguna secara lebih mendalam.

Salah satu algoritma yang sering diterapkan untuk klasifikasi adalah *Support Vector Machine* (SVM). SVM merupakan model yang berbasis teori pembelajaran statistik dan digunakan dalam menyelesaikan tugas-tugas klasifikasi maupun regresi. SVM dikenal karena kemampuannya dalam memisahkan dua kelas data dengan batas keputusan yang optimal. Hal ini didukung oleh penelitian sebelumnya terkait analisis sentimen oleh I. R. Affandi, dkk [4] yang melakukan analisis sentimen terhadap vaksin COVID-19 dengan membandingkan algoritma *Naïve Bayes* dan SVM. Hasil penelitian menunjukkan bahwa algoritma SVM memiliki performa lebih baik dalam hal *accuracy*, *precision*, dan *recall*, dengan nilai masing-masing sebesar 90.47%, 90.23%, dan 90.78%.

Lalu penelitian oleh J. W. Iskandar, dkk [5], mengevaluasi sentimen masyarakat terhadap Samsung Galaxy Z Flip 3 melalui analisis komentar *Youtube* pada Channel *GadgetIn*. Hasil analisis menunjukkan bahwa model klasifikasi SVM memberikan hasil tingkat *accuracy* rata-rata sebesar 96.4%. Lalu penelitian oleh K. A. Rokhman, dkk [6] melakukan analisis sentimen menggunakan data ulasan pada *Google Play* terkait aplikasi transportasi online Gojek. Hasil analisis klasifikasi menunjukkan bahwa metode SVM memiliki tingkat *accuracy* sebesar 90.20%.

Keunggulan SVM terletak pada kemampuannya menangani ruang fitur yang tinggi dan mengatasi ketidakseimbangan data. Dalam dataset kecil SVM

efektif mengatasi masalah *overfitting* yang seringkali muncul saat jumlah data terbatas. Hal ini sejalan dengan penelitian B. W. Sari, dkk [7], yang menggunakan 500 data namun SVM mampu memberikan hasil yang memuaskan, yaitu nilai *accuracy* sebesar 79.6% untuk Telkom dan 83.2% untuk Biznet. Penelitian yang dilakukan oleh B. A. Ardhani, dkk [8], mengenai ulasan tweet tentang kartu prakerja menggunakan dataset sebanyak 500 menunjukkan bahwa pemodelan Support Vector Machine (SVM) dengan jumlah data yang kecil dapat menghasilkan nilai *average accuracy* sebesar 85,20%, nilai *sensitivity* mencapai 91,68%, dan *specificity* sebesar 75,75%. Untuk prediksi sentimen negatif dan positif, masing-masing diperoleh nilai sebesar 85,03% dan 86,04%.

Berdasarkan hasil penelitian sebelumnya, algoritma SVM dapat mengklasifikasikan data dan mendapat hasil akurasi yang baik. Penelitian ini berfokus pada merancang sebuah model SVM yang mampu mengklasifikasikan sentimen masyarakat menjadi kelas positif dan negatif, serta mengukur performa model menggunakan metrik seperti *accuracy*, *precision*, *recall* dan *F1-score*. Selain itu, penelitian ini akan melakukan analisis sentimen dengan mengidentifikasi kata/frasa penting. Hal ini dilakukan dengan memvisualisasikan kata/frasa yang sering muncul dalam ulasan pengguna, menggunakan *wordcloud* untuk mengidentifikasi kata-kata berpengaruh atau yang sering muncul dalam kedua kelas yang dapat memberikan wawasan pandangan pengguna terhadap aplikasi.

Ulasan dari Google Play Store dan App Store merupakan sumber data yang kaya untuk menganalisis pengalaman pengguna. Ulasan ini mencakup rating berbentuk bintang (1-5), yang menggambarkan tingkat kepuasan pengguna terhadap suatu aplikasi. Selain itu, teks ulasan sering kali cukup mendetail, memberikan wawasan tentang kepuasan pengguna, laporan bug, permintaan fitur, dan opini terkait fitur aplikasi, sehingga sangat berguna sebagai masukan untuk pengembangan aplikasi [9]. Dibandingkan dengan sumber lain seperti Twitter, yang memiliki jangkauan lebih luas, data dari media sosial memerlukan proses pembersihan yang lebih rumit dan memakan waktu karena percakapan sering kali tidak fokus pada topik tertentu. Di sisi lain, ulasan di Play Store dan App Store lebih relevan dan terfokus, karena langsung berkaitan dengan aplikasi yang diulas. Hal ini membuat data yang dihasilkan lebih terarah untuk analisis.

Rasa urgensi untuk melakukan penelitian ini timbul dari kebutuhan untuk mengidentifikasi masalah pada aplikasi Whoosh melalui model pembelajaran mesin, berdasarkan kata-kata sentimen negatif yang diidentifikasi. Urgensi ini dipicu oleh pentingnya menjaga kepuasan dan loyalitas pelanggan, yang merupakan

salah satu kunci utama bagi kelangsungan perusahaan [1]. Meningkatkan kualitas aplikasi menjadi langkah krusial dalam upaya ini, dan peningkatan kualitas tersebut dapat dicapai melalui pengembangan atau perbaikan aplikasi.

1.2 Rumusan Masalah

1. Bagaimana implementasi algoritma SVM dapat digunakan untuk menganalisis sentimen ulasan pengguna pada aplikasi Whoosh?
2. Bagaimana performa model dalam hal *accuracy*, *precision*, *recall*, dan *F1-Score* dalam mengklasifikasi sentimen ulasan pengguna pada aplikasi Whoosh?
3. Bagaimana analisis sentimen dapat digunakan untuk mengidentifikasi ulasan sentimen pengguna aplikasi Whoosh?

1.3 Batasan Masalah

1. Sumber data yang diperoleh dari ulasan pada Google Playstore terdapat 349 dan App Store terdapat 206. Data diambil sejak ulasan pertama yakni pada 7 Oktober 2023 hingga ulasan terbaru pada 30 Agustus 2024.
2. penelitian ini terbatas pada penggunaan bahasa Indonesia dalam pengolahan data ulasan.
3. Penelitian analisis sentimen pada aplikasi Whoosh masih terbatas sehingga penelitian ini dapat menjadi sumber informasi baru.

1.4 Tujuan Penelitian

1. Mengimplementasikan algoritma SVM untuk mengklasifikasikan sentimen ulasan pengguna aplikasi Whoosh.
2. Mengukur kinerja model SVM berdasarkan metrik akurasi, presisi, recall, dan *F1-Score* pada ulasan pengguna aplikasi Whoosh..
3. Menentukan sentimen positif dan negatif dari ulasan pengguna aplikasi Whoosh.

1.5 Manfaat Penelitian

- a. Penelitian ini dapat dijadikan referensi bagi penelitian selanjutnya dengan tema terkait agar dapat dikembangkan lebih baik.
- b. Memperoleh informasi tentang efektifitas Algoritma SVM dalam mengklasifikasikan sentimen terhadap aplikasi Whoosh.
- c. Informasi terhadap hasil analisis sentimen dapat menjadi saran untuk membantu pengembangan pada aplikasi Whoosh.

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN

Bab ini mengawali dengan menjelaskan konteks masalah, merumuskan permasalahan yang akan dibahas, menetapkan batasan permasalahan, menjelaskan tujuan dari penelitian ini, serta mengidentifikasi manfaat dari hasil penelitian dan struktur penulisan yang akan diikuti dalam laporan skripsi ini.

- Bab 2 LANDASAN TEORI

Bab ini berfokus pada penjelasan teori-teori yang menjadi dasar bagi penelitian ini. Teori-teori yang dibahas mencakup analisis sentimen, *text mining*, *preprocessing*, dan *Support Vector Machine(SVM)*

- Bab 3 METODOLOGI PENELITIAN

Bab ini menguraikan langkah-langkah penelitian secara sistematis, menyajikan berbagai diagram dan gambar yang menggambarkan perancangan model dalam penelitian ini.

- Bab 4 HASIL DAN DISKUSI

Bab ini mencakup penjelasan tentang hasil implementasi dari penelitian ini. Output yang dihasilkan dari penelitian tersebut akan direpresentasikan dan dianalisis secara mendalam dalam bab ini.

- Bab 5 SIMPULAN DAN SARAN

Bab ini menguraikan rangkuman dari temuan yang diperoleh dari penelitian yang telah dilakukan, serta rekomendasi untuk penelitian yang akan datang.