

BAB I

PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang akan digunakan dalam dokumen skripsi ini.

1.1 Latar Belakang

Access by KAI adalah aplikasi *mobile* yang berfokus pada pemesanan tiket kereta api di Indonesia. Access by KAI melayani pembelian, ubah jadwal, pembatalan, dan transfer tiket berbagai layanan perkeretaapian PT KAI maupun anak perusahaannya. Tidak hanya layanan tiket saja, Access by KAI juga memiliki layanan pendukung untuk menunjang kenyamanan penumpang selama bepergian menggunakan kereta api.

Dalam layanan jasa transportasi publik, opini publik menjadi sangat penting bagi keberlangsungan layanan jasa transportasi publik. Pada saat ini terdapat banyak *platform* yang disediakan PT Kereta Api Indonesia untuk menampung opini publik mengenai kinerja aplikasi Access by KAI diantaranya seperti pada Google Play Store, Apple Store, dan lain-lain. Pengguna yang mengunduh aplikasi Access by KAI melalui Google Play Store dapat memberikan opininya mengenai aplikasi ini melalui fitur Rating dan Ulasan. Pengguna dapat memberikan ulasannya dan juga memberi *rating* kepada aplikasi Access by KAI. *Rating* yang diberikan pengguna dapat memengaruhi *rating* keseluruhan aplikasi di Google Play Store.

Analisis sentimen merupakan salah satu bidang penelitian dalam pemrosesan bahasa alami (*natural language processing*). Analisis sentimen bertujuan untuk memahami opini pengguna terhadap suatu aplikasi dan dapat memberikan wawasan bagi pengembang untuk meningkatkan kualitas layanan. Dalam analisis sentimen fitur yang dihasilkan dan algoritma klasifikasi menjadi poin utama (Atmadja & Purwarianti, 2015). Secara garis besar terdapat dua algoritma pada analisis sentimen, yaitu *rule based method* dan *statistical based method* (Atmadja & Purwarianti, 2015). *Statistical based method* merupakan proses pemisahan data sesuai kelas masing-masing

menggunakan perhitungan matematis atau *machine learning*. Sedangkan *rule based method* menggunakan bantuan *human expert* berupa aturan (*rule*) yang berfungsi untuk memisahkan data sesuai kelas masing-masing.

Sistem analisis sentimen yang dibangun menggunakan algoritma klasifikasi Naïve Bayes. Naïve Bayes merupakan teknik pembelajaran mesin yang bersifat probabilistik. Ciri utama dari algoritma Naïve Bayes adalah asumsi yang naif akan independensi dari masing-masing kondisi atau kejadian (Natalius, 2010). Menurut Gunawan dkk. (2018), kelebihan dari Naïve Bayes adalah proses klasifikasi data dapat disesuaikan dengan sifat dan kebutuhan masing-masing.

Menurut Lailiyah (2017), selain proses klasifikasi, salah satu poin utama dalam analisis sentimen adalah proses ekstraksi fitur. Terdapat dua pendekatan yang bisa digunakan dalam proses ini, yaitu *statistical approach* dan *semantic approach*. Pendekatan statistik menggunakan perhitungan statistik untuk mengekstraksi fitur, seperti frekuensi kemunculan kata dalam dokumen (TF) atau frekuensi kemunculan kata dalam dokumen terhadap keseluruhan dokumen (TF-IDF).

Menurut Lailiyah (2017), pendekatan semantik (makna) menggunakan *lexical resource* untuk mengekstraksi fitur. Analisis sentimen dapat dilakukan melalui pendekatan berbasis leksikon yang membutuhkan pembuatan repositori yang ekstensif, atau pendekatan *machine learning* yang membutuhkan *training* dengan *dataset* yang besar dengan intervensi manual yang minimal.

Dalam konteks ulasan pengguna, *star rating* biasanya digunakan untuk memberikan indikasi cepat mengenai sentimen ulasan dan menjadi standar alat bantu dalam pengambilan keputusan dari ulasan (Al-Natour & Turetken, 2020). Namun *star rating* sering kali tidak cukup mewakili konten ulasan karena *star rating* cenderung sangat subjektif dan bias terhadap kepribadian masing-masing pengulas (Af'idah dkk., 2020). Batasan ini memberikan potensi kepada analisis sentimen, teknik analisis yang menganalisa pola dari teks, alternatif untuk mengidentifikasi sentimen yang sesuai dari sebuah teks, atau untuk menjelaskan ekspresi sebuah teks seperti Star Rating (Al-Natour & Turetken, 2020).

Pada kasus analisis sentimen ulasan pengguna aplikasi Access by KAI, dilakukan pengujian terkait kecocokan opini pengguna terhadap aplikasi dan *rating* bintang yang diberikan. Metode pengujian yang akan dilakukan dengan membandingkan metode *labelling* Part-of-Speech (POS) Tagging dengan metode *labelling* Star Review.

Metode POS Tagging merupakan proses *labelling* kata di dalam sebuah korpus dengan kelas kata yang sesuai berdasarkan definisi dan konteks dari kata tersebut (Maulana & Romadhony, 2021). Dalam pendekatan *rule based*, leksikon digunakan untuk memperoleh POS *tag*, apabila terdapat sebuah kata dengan probabilitas lebih dari satu *tag* maka POS *tag* diperoleh menggunakan aturan yang sesuai. Untuk menghitung polaritas skor dari hasil POS Tagging, akan digunakan metode *rule based*. Khususnya menggunakan *lexicon rule based*. Terdapat dua *library* yang akan digunakan yaitu inSet dan SentiWord.

Pada metode *labelling* Star Review menggunakan TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk membantu dalam mengekstraksi fitur teks ulasan berdasarkan frekuensi kemunculan kata dan pentingnya kata tersebut dalam keseluruhan dokumen. Pada metode ini, Star Rating akan dilabeli menggunakan pendekatan TF-IDF untuk mengidentifikasi kata-kata kunci yang dominan.

Untuk menguji efektivitas dari kedua metode tersebut, analisis sentimen pada ulasan aplikasi Access by KAI dilakukan menggunakan metode Multinomial Naïve Bayes. Metode ini dipilih karena kesederhanaannya dan kemampuannya dalam menangani teks dengan performa yang baik. Kedua metode tersebut nantinya akan dibandingkan untuk menentukan metode yang memiliki kinerja lebih efektif untuk analisis sentimen ulasan pengguna aplikasi Access by KAI.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang dijelaskan, rumusan masalah yang akan diselesaikan dalam penelitian ini adalah bagaimana analisis sentimen pengguna aplikasi Access by KAI menggunakan metode klasifikasi Naïve Bayes.

1.3 Tujuan

Tujuan dilakukan penelitian ini adalah mendapatkan analisis sentimen ulasan pengguna aplikasi Access by KAI dengan membandingkan dua metode *labelling* yaitu POS Tagging dan Star Rating dengan menggunakan metode klasifikasi Naïve Bayes.

1.4 Manfaat

Manfaat dari penelitian ini diharapkan dapat memberikan kontribusi dalam memperluas wawasan tentang efektivitas alat-alat analisis sentimen dalam konteks aplikasi transportasi publik khususnya aplikasi Access by KAI dan dapat memberikan rekomendasi yang lebih tepat kepada pengelola layanan untuk meningkatkan kualitas layanan dan pengalaman pengguna.

1.5 Ruang Lingkup

Ruang lingkup atau batasan masalah dalam penelitian “Analisis Sentimen Opini Pengguna Aplikasi Access by KAI Menggunakan Metode Naïve Bayes Classifier” adalah:

1. Data yang digunakan dalam penelitian ini dibatasi dengan data yang diperoleh dari ulasan aplikasi Access by KAI dalam aplikasi Google Play Store.
2. Data yang digunakan adalah data ulasan pengguna aplikasi Access by KAI sampai tanggal 25 September 2023.
3. Tidak dilakukan pemrosesan terkait *imbalance data* untuk menjaga keaslian data.

1.6 Sistematika Penulisan

Sistematika penulisan dibuat untuk memberikan gambaran umum dan jelas mengenai penulisan serta pembahasan implementasi analisis sentimen opini pengguna aplikasi Access by KAI menggunakan metode klasifikasi Naive Bayes. Berikut adalah sistematika penulisan yang digunakan:

BAB I PENDAHULUAN

Pendahuluan berisikan tentang latar belakang, rumusan masalah, tujuan, manfaat, ruang lingkup, dan sistematika penulisan penelitian analisis sentimen opini pengguna aplikasi Access by KAI menggunakan metode klasifikasi Naive Bayes.

BAB II TINJAUAN PUSTAKA

Bab ini berisi tentang studi pustaka mengenai dasar teori yang berhubungan dengan pelaksanaan dan penyusunan penelitian analisis sentimen opini pengguna aplikasi Access by KAI menggunakan metode klasifikasi Naive Bayes.

BAB III METODOLOGI PENELITIAN

Bab ini berisi langkah-langkah untuk menyelesaikan permasalahan mengenai analisis sentimen opini pengguna aplikasi Access by KAI menggunakan metode klasifikasi Naive Bayes.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari penelitian yang telah dilakukan mengenai analisis sentimen opini pengguna aplikasi Access by KAI menggunakan metode klasifikasi Naive Bayes.

BAB V PENUTUP

Bab ini berisi kesimpulan dari penelitian penulisan analisis sentimen opini pengguna aplikasi Access by KAI menggunakan metode klasifikasi Naive Bayes yang telah dilakukan. Pada bab ini juga terdapat saran untuk penelitian terkait yang dapat dilakukan pada penelitian selanjutnya.