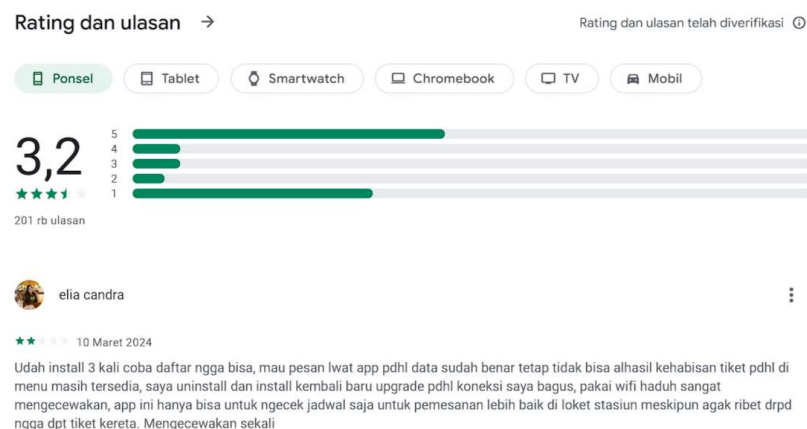


BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era perkembangan teknologi yang semakin maju dan terdepan ini, tidak dapat dipungkiri semua orang bergantung pada teknologi. Dalam berbagai aspek di kehidupan sehari-hari, semua manusia membutuhkan teknologi untuk membantu mempermudah pekerjaan mereka. Salah satunya adalah aktivitas pembelian tiket kereta api. Sekarang untuk membeli tiket kereta api tidak perlu datang ke stasiun, pembeli dapat melakukan pembelian tiket dari rumah. Hal ini dikarenakan pada tahun 2014, PT. Kereta Api Indonesia yang merupakan perusahaan Badan Usaha Milik Negara (BUMN) berinovasi dengan meluncurkan aplikasi KAI Access (Setiawan & Novita, 2021). Tahun 2020, aplikasi ini bahkan memenangkan penghargaan “*the best innovation in payment system*” (bukan bank) dalam Artajasa Award. Fitur dalam KAI Access diantaranya yaitu, KA lokal, KA antar kota, KA bandara, melihat jadwal kereta, melihat kode dan *barcode* pemesanan, dan cek posisi kereta api (Yuda dkk., 2023).



Gambar 1.1. Ulasan dan Rating pada Google Play Store

Gambar 1.1 menunjukkan *Rating* ulasan pada Play Store sebesar 3.2. Pada Juni 2023 jumlah pengguna KAI Access yang tercatat sebanyak 12,4 juta pengguna. Sedangkan untuk pengguna aktif sebanyak 6,1 juta pengguna. Beberapa ulasan dari pengguna yang menyatakan bahwa *loading* aplikasi lambat, batas waktu pemesanan tiket yang terlalu cepat, dan sering terjadi gangguan pada saat pemesanan tiket (Arochma dkk., 2023). Hingga akhirnya pada tanggal 10 Agustus 2023 PT Kereta Api Indonesia melakukan peluncuran aplikasi Access by KAI sebagai bentuk *upgrade* dari aplikasi sebelumnya yaitu KAI Access. Terdapat 4 fitur baru yang ditawarkan pada aplikasi ini antara lain, reservasi hotel, *loyalty poin*, *trip planner*, dan *live tracking*. Dari permasalahan diatas diperlukan klasifikasi ulasan dari pengguna untuk mengetahui sentimen pengguna Access by KAI.

Pada skripsi ini, proses klasifikasi menggunakan data ulasan aplikasi Access by KAI dari Google Play Store dan App Store yang bersifat *open source* yang dapat diakses semua orang. Oleh karena itu, perlu dilakukan proses *scraping* atau pengambilan data berupa ulasan dan *rating* dari pengguna. Kemudian data tersebut akan diproses untuk dianalisis agar mendapatkan nilai akurasi. Data ulasan aplikasi yang akan dikumpulkan dalam skripsi ini adalah data ulasan dari awal *upgrade* aplikasi menjadi Access by KAI di Google Play Store dan App Store pada tanggal 10 Agustus 2023.

Adanya ulasan yang diberikan pengguna untuk aplikasi, perlu dilakukan analisis sentimen untuk melihat bagaimana pendapat dan reaksi pengguna dalam menggunakan aplikasi Access by KAI. *Sentiment analysis*

dapat digunakan untuk menganalisis dokumen dengan terperinci dan merangkum masukan dari pengguna, sehingga pihak pengembang dapat memperoleh hasil analisis tersebut untuk meningkatkan kualitas aplikasi (Kusnadi dkk., 2021). Terdapat dua jenis analisis sentimen yaitu *deep learning models* dan *traditional models*. *Deep learning models* diantaranya yaitu, *Deep Learning Neural Network (DNN)*, *Learning and Teaching Support Network (LTSM)*, *Recurrent Neural Network (RNN)*, dan *Convolutional Neural Network (CNN)*. Sedangkan untuk model tradisional yaitu *Support Vector Machine*, *Naïve Bayes*, dan *Maximum Entropy Classifier* (Ikhsan, 2023).

Analisis sentimen dapat menggunakan beberapa metode klasifikasi. Beberapa metode klasifikasi yang sering digunakan yaitu *Support Vector Machine (SVM)*, *Naïve Bayes*, dan *K-Nearest Neighbor (KNN)*. Pada penelitian yang dilakukan oleh Radiena & Nugroho (2023), menganalisis ulasan KAI Access berbasis aspek menggunakan model CRISP-DM dan metode SVM. Penggunaan SVM dapat digunakan sebagai pilihan dalam analisis sentimen. Terbukti dari hasil pengujian dan nilai rata-rata akurasi yang didapatkan dari semua aspek adalah 92.88%. Sedangkan pada penelitian yang dilakukan oleh Mustakim & Priyanta (2022), menganalisis sentimen ulasan KAI Access berbasis aspek menggunakan algoritma *naïve bayes* dan SVM. Hasilnya adalah tingkat akurasi dari SVM lebih tinggi dibandingkan dengan *naïve bayes*. Kinerja SVM baik dalam memproses data berdimensi tinggi.

Beberapa penelitian terdahulu yang telah menggunakan metode Word2vec untuk ekstraksi fiturnya adalah penelitian dari Suryati dkk. (2023), menggunakan Word2vec dan algoritma SVM untuk menganalisis sentimen

masyarakat terhadap jasa transportasi *online*. Penggunaan Word2vec menggunakan model skip-gram dinilai cukup baik untuk analisis sentimen dilihat dari hasil evaluasi performa model. Penelitian lainnya oleh Purnasiwi dkk. (2023), menggunakan *word embedding* Word2vec dan metode LSTM untuk menganalisis sentimen ulasan produk *skincare*. Penerapan Word2vec dapat dikatakan mampu menunjukkan makna dari informasi yang terbatas dalam teks yang dikumpulkan dengan baik. Sedangkan penelitian yang dilakukan oleh Utama & Masruro (2022), menggunakan metode Word2vec dan algoritma SVM untuk analisis sentimen *tweet* penerbangan di Amerika pada Februari 2015. Hasil dari metode Cross Validation dengan fold 5 adalah nilai akurasi SVM Linear sebesar 75,20%.

Berdasarkan penelitian terdahulu yang telah dijabarkan diatas mengenai analisis sentimen dan dengan mempertimbangkan hasil analisisnya, maka skripsi ini akan menggunakan algoritma *Support Vector Machine* (SVM) sebagai metode klasifikasi sentimen dan metode Word2vec untuk ekstraksi fitur. Setelah proses pembuatan model selesai, diperlukan evaluasi performa model dengan menggunakan *confusion matrix* untuk mengukur tingkat akurasi dari model yang dipakai sebelumnya. Pada tahap inilah tingkat kepuasan pengguna dalam menggunakan aplikasi Access by KAI dapat diketahui. Dengan adanya skripsi ini diharapkan dapat menghasilkan nilai akurasi dari klasifikasi sentimen menggunakan algoritma *Support Vector Machine* (SVM) dan mengetahui pengaruh metode Word2vec terhadap performa model klasifikasi.

1.2 Rumusan Masalah

Adapun rumusan masalah berdasarkan latar belakang dari skripsi ini yaitu “Bagaimana mengevaluasi dan menerapkan algoritma *Support Vector Machine* (SVM) dan metode Word2vec terhadap performa model klasifikasi sentimen ulasan aplikasi Access by KAI?”

1.3 Batasan Masalah

Pada skripsi ini terdapat batasan masalah untuk membatasi ruang lingkup yang akan dibahas pada skripsi ini, yaitu sebagai berikut:

1. Data ulasan pengguna yang digunakan adalah ulasan dari tanggal 10 Agustus 2023 sampai dengan 5 Februari 2024.
2. Data ulasan Access by KAI yang diambil dari Google Play Store dan App Store yang berbahasa Indonesia.
3. Pelabelan data dibagi menjadi dua kelas yaitu positif dan negatif.

1.4 Tujuan

Adapun tujuan dari skripsi ini berdasarkan latar belakang dan rumusan masalah yang telah dijelaskan adalah mengevaluasi dan menerapkan algoritma *Support Vector Machine* (SVM) dan metode Word2vec terhadap performa model klasifikasi sentimen ulasan aplikasi Access by KAI.

1.5 Sistematika Penulisan

Skripsi ini memiliki sistematika penulisan mencakup lima bab sebagai berikut:

BAB I PENDAHULUAN

Pada bab pertama ini membahas beberapa hal diantaranya, latar belakang dari skripsi ini, batasan masalah, tujuan, dan sistematika penulisan skripsi ini.

BAB II TINJAUAN PUSTAKA

Pada bab kedua ini membahas dasar teori yang digunakan dalam skripsi ini yaitu mengenai analisis sentimen. Selain itu, bab ini juga membahas penelitian terdahulu sebagai pembanding dengan skripsi ini.

BAB III METODOLOGI PENELITIAN

Pada bab ketiga ini berisi pembahasan mengenai tahapan dari metode SEMMA yang digunakan pada skripsi ini. Tahapannya berupa *sample, explore, modify, model, dan access*.

BAB IV HASIL DAN PEMBAHASAN

Pada bab keempat ini mencakup hasil dan pembahasan dari setiap tahapan penelitian yang dilakukan.

BAB V PENUTUP

Pada bab kelima ini memuat kesimpulan dari hasil analisis sentimen yang telah dilakukan dan juga saran untuk mengembangkan penelitian selanjutnya.