

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan Teknologi yang pesat saat ini mempengaruhi berbagai aspek kehidupan menjadi serba mudah dan efisien, seiring dengan penetrasi teknologi digital yang sangat dalam dan digunakan secara luas, dampak teknologi digital akan semakin terasa, terutama di dunia transportasi online dan dunia startup digital. Salah satunya perusahaan unicorn di Indonesia, Gojek telah berkembang menjadi platform super app yang tidak hanya menyediakan layanan transportasi, tetapi juga layanan lainnya Pembayaran digital, Pengiriman makanan, Pengiriman barang, Layanan pembelian dan pengantaran, Layanan kesehatan [1].

Gojek Indonesia atau yang memiliki nama lain PT Aplikasi Karya Anak Bangsa satu ini merupakan layanan pemesanan ojek melalui aplikasi mobile. Aplikasi ini dapat di unduh di Apple Store maupun Play Store dan merupakan karya putra bangsa bernama, Nadiem Makarim [2]. Gojek adalah perusahaan berjiwa sosial yang memimpin revolusi industri transportasi ojek. Gojek bermitra dengan para pengendara ojek, Gojek menjadi solusi utama dalam mengirimkan barang, pesan antar makanan, berbelanja dan berpergian ditengah kemacetan. Gojek adalah Aplikasi karya anak Bangsa satu ini merupakan layanan pemesanan ojek melalui aplikasi mobile [2].

Seperti yang telah dipaparkan di atas, Gojek Indonesia dicetuskan oleh seorang WNI yang bernama Nadiem Makarim yang merupakan lulusan program Master of Business Administration di sekolah bisnis Harvard. Dia terinspirasi mendirikan perusahaan Gojek Indonesia dari pengalamannya yang harus setiap hari menembus kemacetan di Jakarta dengan menggunakan transportasi ojek. Pada saat itu Nadiem Makarim masih berstatus sebagai Managing Editor Zalora Indonesia sekaligus Chief Innovation Officer Kartuku [2].

Analisis sentiment adalah sebuah proses untuk menentukan sentiment atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bias dikategorikan sebagai sentiment positif atau negatif [3]. Analisis ulasan secara mudah dapat dilakukan dengan melihat jumlah bintang yang diberikan oleh pembeli, tetapi jumlah bintang tidak dapat mewakili isi dari keseluruhan ulasan. Diperlukan melihat seluruh isi komentar ulasan untuk dapat mengetahui keseluruhan maksud

ulasan. Sangat dimungkinkan untuk menganalisis ulasan secara manual dengan melihat satu persatu, namun apabila ulasan yang dimiliki banyak akan lebih cepat menggunakan sistem analisis sentimen [4].

NLP memungkinkan komputer untuk memahami, menafsirkan, dan memanipulasi bahasa manusia, sedangkan machine learning memungkinkan sistem untuk belajar dari data dan meningkatkan kinerjanya dari waktu ke waktu [5]. Tahapan NLP untuk memproses dan menganalisis mencakup pengumpulan data ulasan pengguna, preprocessing data, pemilihan dan penerapan algoritma machine learning serta evaluasi kinerja model yang dihasilkan [5]. Analisis Sentimen dapat digunakan sebagai alat untuk penelitian terkait, karena dapat membantu mengidentifikasi dan menganalisis perasaan dan opini yang terkait dengan topik tersebut di media sosial [5]. Analisis sentimen dalam ulasan pengguna aplikasi gojek di play store ini menyajikan hasilnya melalui kombinasi analisis kuantitatif (metrik evaluasi, statistik) dan kualitatif (contoh ulasan, tema umum). Visualisasi yang jelas dan ringkas yang terstruktur memudahkan pembaca untuk memahami temuan penelitian dan implikasinya. Dan memberikan rekomendasi bagi pengembang aplikasi Gojek berdasarkan penelitian analisis sentimen ini. Dengan pendekatan ini, penelitian ini tidak hanya memberikan wawasan tentang sentimen pengguna Gojek tetapi juga menunjukkan efektivitas model dalam tugas analisis sentimen. Dalam ekosistem digital, ulasan pengguna di platform seperti Google Play Store merupakan aset kritis yang mencerminkan pengalaman langsung pengguna. Ulasan ini menjadi sumber umpan balik yang berharga untuk mengukur kepuasan, mengidentifikasi masalah, dan merancang strategi pengembangan produk. Namun, volume ulasan yang sangat besar mencapai ribuan bahkan ratusan ribu menjadikan analisis manual tidak lagi feasible, lambat, dan rentan terhadap subjektivitas. Selain itu, ditemukan inkonsistensi antara rating bintang dan isi teks ulasan, seperti komentar bernada positif yang diberi bintang satu, atau sebaliknya. Inkonsistensi ini mengaburkan pemahaman sebenarnya dari sentimen pengguna dan dapat menyesatkan dalam pengambilan keputusan.

Berbagai metode klasifikasi telah dikembangkan untuk meningkatkan akurasi dalam analisis sentimen, di antaranya adalah Naïve Bayes dan SVM. Naïve Bayes adalah metode berbasis probabilitas yang cepat dan efisien untuk klasifikasi teks[6]. Sedangkan SVM merupakan algoritma berbasis vektor yang bekerja

dengan mencari hyperplane optimal untuk memisahkan data ke dalam kategori tertentu[6].

Sebagai contoh penelitian-penelitian sebelumnya seperti penelitian yang ditulis di jurnal yang berjudul Perbandingan Algoritma Support Vector Machine (SVM) dan Naive Bayes pada Analisis Sentimen Bahasa Jawa dan Sunda. Galuh Kusuma Putria, Herry Sujainia, Desepta Isna Ulumia. Vol. 13, No. 2, April 2025, yang menjelaskan setiap algoritma memiliki keunggulan tersendiri dalam hal kecepatan dan akurasi. Algoritma Support Vector Machine (SVM) sangat baik untuk menangani data berukuran besar dan memberikan hasil yang konsisten. Sebaliknya, Naive Bayes (NB) adalah algoritma yang lebih sederhana namun memiliki keunggulan dalam kecepatan proses [7]. Dan kesimpulan dari hasil penelitian tersebut bahwasannya Dapat diambil kesimpulan bahwa algoritma SVM terbukti lebih unggul dibandingkan dengan Naive Bayes dalam klasifikasi sentimen teks berbahasa Jawa dan Sunda, dengan akurasi tertinggi yaitu di Bahasa Jawa 76% dan Bahasa Sunda 80%. SVM juga memberikan hasil yang lebih baik dalam hal precision, recall, dan f1-score. Dalam hal klasifikasi teks bahasa daerah Jawa dan Sunda [7].

Berdasarkan penjelasan di atas, penelitian ini bertujuan untuk melakukan analisis sentimen pada ulasan pengguna gojek untuk menerapkan teknik NLP, khususnya model TF-IDF, dalam menganalisis sentimen ulasan pengguna Gojek di Play Store. Tujuannya adalah memudahkan mengklasifikasikan ulasan yang jumlahnya mencapai ribuan ke dalam kategori sentimen positif dan negatif serta mengevaluasi performa model TF-IDF dalam tugas klasifikasi sentimen. Dan juga memberikan informasi yang dibutuhkan bagi pengolah aplikasi gojek dalam pengembangan lebih lanjut.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Dalam konteks latar belakang mengenai aplikasi gojek dan ulasan dari penggunanya, gojek telah menerima banyak tanggapan atau sentimen dari masyarakat, terutama melalui ulasan di play store. Oleh karena itu, analisis sentimen sangat diperlukan untuk menentukan apakah mayoritas tanggapan tersebut positif atau negatif. penelitian ini bertujuan untuk melakukan analisis sentimen pada ulasan pengguna gojek untuk menerapkan NLP, khususnya model

TF-IDF, dalam menganalisis sentimen ulasan pengguna Gojek di Play Store. Tujuannya adalah mengklasifikasikan ulasan ke dalam kategori sentimen positif dan negatif serta mengevaluasi performa model TF-IDF dalam tugas klasifikasi sentimen. Dan juga memberikan informasi yang dibutuhkan bagi pengolah aplikasi gojek dalam pengembangan lebih lanjut.

1.2.2 Metode Usulan

Penelitian ini dilakukan dalam beberapa tahap. Pertama, data ulasan pengguna aplikasi gojek dikumpulkan dari Play Store menggunakan teknik web scraping. Data kemudian diproses melalui tahapan preprocessing, termasuk pembersihan teks, tokenisasi, normalisasi, dan stopwords. Selanjutnya, model TF-IDF melakukan tugas klasifikasi sentimen dengan melakukan ekstraksi fitur menggunakan dataset ulasan yang telah diproses. Pembagian dataset dilakukan dengan rasio 80:20, 90:10, dan 70:30 untuk data latih dan data uji. Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score.

1.2.3 Pertanyaan Penelitian

1. Bagaimana performa svm pada analisis sentiment ulasan pengguna gojek di playstore?
2. Bagaimana pengaruh perubahan kernel dan komposisi split dataset terhadap sistem analisis sentimen ulasan pengguna gojek menggunakan metode SVM?

1.3 Tujuan Dan Manfaat

1.3.1 Tujuan Penelitian

1. Mengimplementasikan model TF-IDF untuk analisis sentimen pada ulasan pengguna aplikasi gojek di play store.
2. Membangun model yang dapat mengklasifikasikan ulasan pengguna ke dalam kategori sentimen positif dan negatif secara otomatis.
3. Mengevaluasi performa model dalam tugas analisis sentimen menggunakan metrik seperti akurasi, precision, recall, dan F1-score.

1.3.2 Manfaat Penelitian

Dalam penelitian yang sekarang sedang dilakukan, hasil dari analisis sentimen ini dilihat dari manfaat praktisnya untuk pengembang gojek dapat di jadikan sebuah sumber informasi yang dibutuhkan bagi pengembang aplikasi gojek dalam pengembangan yang lebih baik lagi. Pada penelitian ini analisis sentimen ulasan pengguna gojek yang dilakukan dengan menerapkan Support Vector Machine (SVM) dengan pendekatan model TF-IDF. Aplikasi ini nantinya akan dapat memberikan sebuah informasi ulasan atau opini yang dapat mengartikan sebuah opini tersebut apakah termasuk ke kategori positif atau negatif.

1.4 Batasan-Batasan

1. Data ulasan tersebut diambil dari aplikasi gojek di Play Store melalui teknik scraping
2. Data ini berisikan ulasan dalam bahasa Indonesia dan tidak termasuk ulasan dalam bahasa lain.
3. Pembagian kategori atau label di bagi menjadi 2 bagian yaitu positif dan negatif.
4. Tugas akhir ini menfokuskan pada setiap ulasan tersebut pada kategori di setiap ulasan.

1.5 Sintematika Penulisan Proposal

Pada penelitian ini rencana susunan sistematika penulisan proposal sebagai berikut:

BAB I PENDAHULUAN

Bab ini memuat latar belakang penelitian, rumusan masalah penelitian klasifikasi teks, batasan masalah penelitian, tujuan penelitian, keunggulan penelitian, dan sistem penulisan.

BAB II KAJIAN PUSTAKA

Pada bab ini dilakukan kajian terhadap literatur dan teori terkait penelitian yang dilakukan yang diperoleh dari scrapping pada aplikasi gojek di play store. Beberapa teori yang digunakan dalam penelitian ini antara lain natural language processing (NLP), analisis Sentimen, pengumpulan data, scraper data, preprocessing data, TF-IDF, support vector machine, dan confusion matrix dalam analisis sentimen ulasan pengguna aplikasi gojek di play store.

BAB III METODE USULAN

Dalam bab ini dirujuk langkah-langkah penelitian agar penelitian yang dilakukan mencapai tujuan yang maksimal. Tahapan penelitian meliputi tahap penelitian, perancangan sistem, pengumpulan data, tahap preprocessing, tahap klasifikasi, tahap evaluasi dan tahap skenario pengujian.

BAB IV HASIL DAN PEMBAHASAN

Bab IV membahas hasil analisis terhadap data yang telah diproses, serta penerapan metode yang diusulkan. Penulis akan mengevaluasi kinerja model SVM menggunakan teknik TF-IDF. Analisis dilakukan berdasarkan metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-Score*.

BAB V KESIMPULAN

Bab V menyimpulkan hasil penelitian, menjawab rumusan masalah, serta memberikan gambaran mengenai efektivitas SVM dalam analisis sentimen. Saran diberikan untuk pengembangan model ke depan dengan fokus pada peningkatan layanan berdasarkan analisis sentimen pengguna.