

BAB I PENDAHULUAN

1.1 Latar Belakang

Gojek merupakan aplikasi ojek online yang didirikan oleh Nadiem Makarim pada tahun 2010[1]. Berdasarkan data dari Databoks, Gojek berada pada urutan pertama pada kategori aplikasi transportasi online terbanyak di unduh di RI 2023, dengan rata-rata 957 ribu unduhan perbulan dari pengguna *smartphone*[2]. Meskipun demikian, hal tersebut tidak bisa dijadikan sebagai tolak ukur untuk menilai kepuasan pengguna. Kepuasan pengguna hanya dapat dinilai dari rating dan komentar yang diberikan oleh pengguna Gojek di Google Play Store.

Pengguna aplikasi Gojek memiliki berbagai cara untuk menyampaikan opini mereka, salah satunya melalui kolom komentar yang disediakan di Google Play Store. Meskipun tidak semua pengguna menyalurkan masukan atau keluhan di sana, beberapa mungkin memilih langsung menghubungi layanan pelanggan. Dan kolom komentar tetap menjadi sumber data yang berharga. Komentar yang ditulis mencakup beragam opini, baik positif maupun negatif, sehingga memerlukan analisis yang sistematis untuk memahami sentimen yang terkandung. Penelitian ini bertujuan untuk menerapkan analisis sentimen guna mengkategorikan komentar pengguna aplikasi Gojek di Google Play Store menjadi komentar positif dan negatif[3].

Menurut Bing Liu analisis sentimen merupakan proses yang digunakan untuk menentukan opini atau sikap seseorang terhadap suatu topik atau produk melalui analisis teks. Tujuan dari analisis ini adalah untuk mengidentifikasi apakah sentimen yang terkandung dalam teks tersebut bersifat positif, negatif, atau netral. Revisi sumber data yang valid[4].

Menurut penelitian yang dilakukan Anggi Puji Astuti, Syariful Alam dan Irsan Jaelani dengan judul penelitian "Komparasi Algoritma SVM dengan *Naïve Bayes* Untuk Analisis Sentimen Pada Aplikasi BRImo". Hasil penelitian ini menunjukkan bahwa algoritma SVM lebih baik dalam melakukan klasifikasi yaitu dengan akurasi 97,69%, sedangkan algoritma *Naïve Bayes* mendapatkan akurasi sebesar 96,53%[5]. Sedangkan pada penelitian yang dilakukan oleh Arman

Ardiansyah dan Wahyudin yang berjudul “Analisis Sentimen Pada Ulasan Aplikasi Home Credit dengan Metode SVM dan KNN”. Hasil dari penelitian ini menunjukkan bahwa algoritma SVM lebih baik dalam melakukan klasifikasi dengan akurasi sebesar 88%, sedangkan algoritma KNN mendapatkan akurasi sebesar 79%[6].

Berdasarkan beberapa referensi yang telah disebutkan, algoritma SVM terbukti efektif dalam melakukan analisis sentimen dibandingkan dengan algoritma lainnya. Oleh karena itu, penelitian ini memilih algoritma SVM sebagai metode klasifikasi untuk analisis sentimen. Meskipun demikian, belum ada hasil pasti mengenai kernel yang paling akurat dalam melakukan analisis sentimen. Oleh karena itu, penelitian ini akan membandingkan akurasi antara kernel *Linear*, kernel RBF, kernel *Sigmoid*, dan kernel *Polynomial* dalam analisis sentiment menggunakan dua metode pelabelan, yaitu metode pelabelan *TextBlob* dan pelabelan manual. Dikarenakan beberapa referensi sebelumnya tidak menggunakan metode untuk mengatasi ketidakseimbangan data, penelitian ini akan menggunakan teknik SMOTE untuk menyeimbangkan data dan mengamati bagaimana hal ini dapat mempengaruhi akurasi dari kernel SVM. Temuan dari penelitian ini dapat menjadi panduan bagi pihak manajemen Gojek untuk mengembangkan aplikasinya dan meningkatkan kualitas layanan.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, ada beberapa rumusan masalah sebagai berikut:

1. Berapa hasil akurasi dari kernel *Linear*, kernel RBF, kernel *Sigmoid*, dan kernel *Polynomial* pada pelabelan manual?
2. Berapa hasil akurasi dari kernel *Linear*, kernel RBF, kernel *Sigmoid*, dan kernel *Polynomial* pada pelabelan *TextBlob*?

1.3 Batasan Masalah

Berdasarkan rumusan masalah di atas, maka permasalahan yang akan di batasi adalah sebagai berikut;

1. Penelitian ini menggunakan data komentar dan rating yang diberikan

oleh pengguna aplikasi Gojek pada Google play store.

2. Penelitian ini akan membandingkan kinerja dari kernel *Linear*, kernel RBF, kernel *Sigmoid*, dan kernel *Polynomial* menggunakan metode pelabelan TextBlob dan pelabelan manual dalam analisis sentimen komentar Gojek di Google Play Store.
3. Evaluasi performa dari kernel *Linear*, kernel RBF, kernel *Sigmoid*, dan kernel *Polynomial* akan dilakukan menggunakan *confusion matrix*.
4. Penelitian hanya menggunakan dua sentimen yaitu, sentimen positif dan negatif.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai dalam penelitian ini adalah mengevaluasi kinerja kernel *Linear*, kernel RBF, kernel *Sigmoid*, dan kernel *Polynomial*. menggunakan metode pelabelan TextBlob dan pelabelan manual, yang bertujuan untuk menentukan kernel mana yang menghasilkan akurasi terbaik. Selain itu, penelitian ini bertujuan untuk melihat opini pengguna tentang aplikasi Gojek dari Google play Store.

1.5 Manfaat Penelitian

Manfaat yang diinginkan dari penelitian ini adalah sebagai berikut:

1. **Manfaat Teoritis**
 Penelitian ini menghasilkan pengetahuan yang lebih luas tentang bagaimana kernel-kernel *Support Vector Machine* berkeja dalam kasus ini.
2. **Manfaat Praktis bagi Objek Penelitian**
 Hasil penelitian ini dapat menghasilkan informasi bagi pengembang aplikasi Gojek untuk memperbaiki kualitas pelayan dalam mengembangkan sistem pada aplikasi gojek.
3. **Manfaat bagi Peneliti Selanjutnya**
 Hasil dari penelitian ini seperti dataset dan metodologi yang dipakai. Dapat sebagai sumber referensi bagi peneliti yang ingin melakukan penelitian dalam masalah yang sama.

1.6 Sistematika Penulisan

Penyusunan laporan skripsi akan menggunakan sistematika penulisan yang akan terbagi menjadi beberapa bab sebagai berikut:

BAB I PENDAHULUAN

Berisi Latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab II Tinjauan pustaka ini mencakup :

- A. Study literatur, berisi tentang penelitian sebelumnya yang ada kaitannya dengan penelitian ini.
- B. Dasar teori, berisi penjelasan tentang konsep dasar algoritma SVM dan kernel-kernel yang di pakai dalam penelitian. Hal ini mencakup bagaimana cara kerja kelebihan dan kelemahan algoritma SVM. Kemudian dasar teori juga berisi tentang analisis sentimen yang menjelaskan tentang definisi, tujuan dan metode yang digunakan dalam klasifikasi sentimen. Ini mencakup teknik-teknik analisis teks seperti data *cleaning*, *case folding*, *stopword removal*, *tokenizing*, *stemming*, TF-IDF, SMOTE dan klasifikasi sentimen. Dasar teori juga berisi tentang Google Play Store yang digunakan sebagai latar belakang penelitian. Dan yang terakhir dasar teori berisi tentang *confusion matrix* yang dipakai untuk mengevaluasi kinerja algoritma. Ini mencakup *accuracy*, *presisi*, *recall*, dan *f1-Score*.

BAB III METODE PENELITIAN

Bab III akan membahas alur penelitian yang diterapkan dalam penelitian ini. Alur penelitian ini mencakup pengumpulan data, pelabelan, *preprocessing*, *exploratory data analysis*, visualisasi *word clouds*, ekstraksi fitur, SMOTE, klasifikasi dan evaluasi hasil klasifikasi

BAB IV HASIL DAN PEMBAHASAN

Bab IV akan membahas hasil dari penelitian yang sudah dilakukan. Ada beberapa aspek yang akan dijelaskan di antaranya adalah ada berapa banyak komentar yang diklasifikasikan sebagai sentimen positif dan negatif. Selain itu, penulis juga akan membahas hasil kinerja dan pengaruh dari penggunaan SMOTE,

menggunakan kernel SVM yaitu kernel *Linear* dan kernel *Radial Basis Function* (RBF), disini juga akan dijabarkan hasil evaluasi menggunakan *confusion matrix* yang mencakup *accuracy*, *precision*, *recall* dan *f1-score*. Selanjutnya, penulis akan membahas tentang kata kata yang sering muncul dalam visualisasi *word clouds*.

BAB V PENUTUP

Bab V akan mencakup tentang kesimpulan dan saran. Kesimpulan dalam penelitian ini digunakan untuk menyajikan ringkasan utama atau pokok dari penelitian yang telah dilaksanakan, dan saran digunakan untuk dapat membantu peneliti dimasa depan dalam membuat penelitian yang lebih relevan.

