

## ***Penggunaan Algoritma Support Vector Machine (SVM) untuk Menganalisis Sentimen dari Ulasan Pelanggan Terhadap Layanan Kurir J&T Express di Google Play Store***

Saddam Mahendra<sup>1</sup>, Suzuki Syofian<sup>2</sup>

<sup>1</sup>Program studi Teknologi Informasi, Universitas Darma Persada

<sup>2</sup>Dosen Program Studi Teknologi Informasi, Universitas Darma Persada,

Jl. Taman Malaka Selatan No.22, Pondok Kelapa, Duren Sawit, DKI Jakarta, Indonesia 13450

\*Koresponden : [suzukiumb@gmail.com](mailto:suzukiumb@gmail.com)

### **Abstrak**

Kurir adalah salah satu layanan yang sering dimanfaatkan oleh masyarakat, terutama di era digital saat ini. Layanan ini memungkinkan pengirim untuk mengirim barang atau dokumen tanpa harus pergi langsung ke tujuan. Di Indonesia, J&T Express menjadi favorit banyak orang sebagai penyedia layanan pengiriman. Walaupun persaingan semakin ketat dan teknologi terus berkembang, kualitas layanan kurir tetap menjadi hal yang penting bagi pelanggan. Namun, perlu dicatat bahwa aplikasi J&T Express di Google Play Store mendapat rating rendah, yang menjadi latar belakang dari penelitian ini. Penelitian ini bertujuan untuk mengukur tingkat kepuasan pelanggan terhadap layanan kurir J&T Express melalui analisis ulasan di Google Play Store. Metode yang digunakan adalah analisis sentimen dengan menggunakan algoritma Support Vector Machine, dengan proses yang mengikuti metodologi CRISP-DM. Hasil penelitian menunjukkan bahwa dari pemahaman bisnis hingga pemodelan, kinerja Support Vector Machine menunjukkan hasil yang memuaskan. Selain itu, penelitian ini juga menghasilkan implementasi yang dapat diakses melalui situs web di alamat [jnt-sentiment.streamlit.app](http://jnt-sentiment.streamlit.app). Harapannya, penelitian ini dapat membantu J&T Express untuk memahami lebih dalam kepuasan pelanggan mereka dan meningkatkan kualitas layanan mereka secara keseluruhan.

**Kata kunci** — Analisis Sentimen, Crisp-Dm, Support Vector Machine, J&T Express, Google Play Store

### **Abstract**

Couriers are a service that is often used by the public, especially in the current digital era. This service allows senders to send goods or documents without having to go directly to the destination. In Indonesia, J&T Express is many people's favorite as a delivery service provider. Even though competition is getting tougher and technology continues to develop, the quality of courier service remains important for customers. However, it should be noted that the J&T Express application on the Google Play Store received a low rating, which is the background for this research. This research aims to measure the level of customer satisfaction with J&T Express courier services through analysis of reviews on the Google Play Store. The method used is sentiment analysis using the Support Vector Machine algorithm, with a process that follows the CRISP-DM methodology. The research results show that from business understanding to modeling, the performance of the Support Vector Machine shows satisfactory results. Apart from that, this research also produced an implementation that can be accessed via the website at [jnt-sentiment.streamlit.app](http://jnt-sentiment.streamlit.app). The hope is that this research can help J&T Express to understand more deeply their customer satisfaction and improve their overall service quality. .

**Keyword** : analisis sentimen, crisp-dm, support vector machine, j&t express, google play store.

## **1. PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

Beberapa dekade terakhir, kemajuan teknologi informasi telah memfasilitasi pengolahan data dengan lebih canggih dan efisien. Dalam konteks ini, teknologi data mining menjadi sangat penting untuk membantu organisasi mendapatkan informasi yang lebih mendalam dan akurat dari data yang tersimpan. Dengan kemampuan untuk mengelola data yang lebih besar dan lebih kompleks, teknologi data mining memainkan peran yang semakin penting dalam proses pengambilan keputusan.

Analisis sentimen, bagian dari Pemrosesan Bahasa Alami (Natural Language Processing/NLP), berkembang untuk mengidentifikasi dan mengekstrak pendapat dalam bentuk teks. Saat ini, informasi dalam bentuk teks banyak tersedia di internet melalui forum, blog, media sosial, dan situs yang berisi ulasan. Dengan

bantuan analisis sentimen, informasi yang awalnya tidak terstruktur dapat diubah menjadi data yang lebih terstruktur.

Pelayanan kurir merupakan salah satu layanan pengiriman barang atau dokumen yang banyak dimanfaatkan, terutama di era digital ini. Dengan layanan kurir, pengirim dapat menggunakan jasa pengiriman yang telah tersedia tanpa perlu datang langsung ke tujuan. J&T Express adalah salah satu dari banyaknya ekspedisi pengiriman di Indonesia yang menyediakan layanan pengiriman barang atau produk ke seluruh wilayah Indonesia, memberikan kemudahan bagi masyarakat untuk memenuhi kebutuhan mereka.

Namun, dengan kemajuan teknologi dan persaingan yang semakin ketat, kualitas layanan kurir menjadi sangat penting bagi pelanggan. Keterlambatan pengiriman, kerusakan barang, dan pelayanan yang kurang memuaskan dapat berdampak negatif terhadap kepercayaan dan reputasi layanan kurir tersebut.

Aplikasi J&T Express di Google Play Store mendapat rating rendah, itulah sebabnya penelitian ini dilakukan. Fokus penelitian ini adalah untuk mengeksplorasi tingkat kepuasan terhadap layanan kurir yang tercermin dalam ulasan Google Play Store, dengan tujuan mengidentifikasi masalah yang mungkin timbul pada layanan kurir J&T Express. Selain itu, penelitian ini akan menggunakan analisis sentimen dengan algoritma Support Vector Machine (SVM) untuk mengukur kepuasan pelanggan terhadap layanan kurir J&T Express. Data untuk penelitian ini diambil dari hasil web scraping Google Play Store yang berisi ulasan pengguna aplikasi J&T Express. Harapannya, penelitian ini dapat memberikan kontribusi bagi J&T Express untuk memahami pandangan pelanggan dan meningkatkan kualitas layanan yang mereka tawarkan.

Berdasarkan penjelasan di atas, rumusan masalah penelitian ini adalah:

1. Berapakah tingkat akurasi yang dihasilkan oleh metode SVM dalam mengklasifikasikan tingkat kepuasan pelanggan di J&T Express?
2. Bagaimana kinerja metode Support Vector Machine (SVM) dalam mengklasifikasikan tingkat kepuasan pelanggan di J&T Express?

Tujuan penelitian ini adalah untuk mengetahui tingkat kepuasan pelanggan terhadap layanan kurir J&T Express dengan menggunakan algoritma SVM, menilai akurasi algoritma SVM, mengevaluasi performa algoritma SVM, dan menerapkan analisis sentimen J&T Express pada platform web yang sudah dilatih. Manfaat dari penelitian ini diharapkan dapat memberikan informasi yang berharga tentang kepuasan pelanggan terhadap layanan kurir J&T Express.

## **2. LANDASAN TEORI**

### **2.1 Analisis *Sentiment***

Menurut penelitian yang dilakukan oleh Pasek et al. pada tahun 2022, analisis sentimen merujuk pada proses klasifikasi emosi dalam tulisan menjadi kategori positif, negatif, atau netral. Analisis sentimen juga dikenal sebagai Opinion Mining, suatu bidang yang bertujuan untuk menganalisis pendapat, evaluasi sentimen, sikap, penilaian, dan emosi seseorang dalam teks atau bahasa tertulis. Dari penjelasan ini, dapat disimpulkan bahwa analisis sentimen melibatkan identifikasi perasaan atau respons seseorang yang diungkapkan melalui teks, yang kemudian dikelompokkan menjadi sentimen netral, positif, atau negatif.

### **2.2 *Data Mining***

Menurut hasil penelitian yang dilakukan oleh Pratama dan rekan pada tahun 2022, Data Mining dijelaskan sebagai suatu proses penemuan hubungan, pola, dan tren yang berharga dengan cara menyelidiki sejumlah besar data dari berbagai sumber, menggunakan teknologi pengenalan pola seperti statistik dan teknik matematika. Data Mining merupakan kegiatan analisis data yang melibatkan penggunaan perangkat lunak untuk mengidentifikasi pola dan aturan yang ada dalam kumpulan data. Pendekatan ini dimanfaatkan untuk mengungkap pengetahuan yang tersembunyi dalam kumpulan data yang besar.

### **2.3 *Text Mining***

Berdasarkan penelitian yang dilakukan oleh Ramadhani & Wahyudin pada tahun 2022, Text Mining diterangkan sebagai suatu tahapan dalam domain Data Mining yang merujuk pada ekstraksi informasi dari teks atau data yang umumnya terdapat dalam dokumen. Fokus Text Mining adalah untuk mengungkap wawasan baru atau frasa khusus, yang memerlukan analisis yang berhubungan erat antara dokumen satu dengan yang lain.

### **2.4 *CRISP-DM***

Menurut riset yang dilakukan oleh Kurniawan & Yasir pada tahun 2022, Cross Industry Standard Process for Data Mining (CRISP-DM) dijelaskan sebagai sebuah kerangka kerja proses yang digunakan secara meluas

untuk keperluan Data Mining, tidak bergantung pada sektor industri tertentu. Model ini terdiri dari enam tahapan siklus yang berulang, dimulai dari Pemahaman Bisnis, Pemahaman Data, Persiapan Data, Pemodelan, Evaluasi, hingga Implementasi. CRISP-DM dianggap sebagai metode yang sangat komprehensif dalam konteks pengembangan proyek industri, dan telah menjadi yang paling dominan dalam proyek-proyek analitik, Data Mining, serta studi Data.

### 2.5 Preprocessing

Penelitian yang dilakukan oleh Gifari et al. pada tahun 2022 menunjukkan bahwa preprocessing memiliki tujuan untuk menormalkan istilah dalam kalimat, sehingga data latih dan fitur yang diekstrak sesuai dengan kebutuhan. Proses ini bertujuan untuk mempermudah pengolahan data. Terdapat enam tahap dalam pra-pengolahan, yaitu:

1. Normalisasi: Melibatkan pemastian bahwa kata-kata yang diperpanjang atau disingkat diubah menjadi kata-kata yang normal menurut Kamus Besar Bahasa Indonesia (KBBI). Konversi slang digunakan untuk mengubah kata tidak baku menjadi kata baku.
2. Pembersihan: Aksara non-alfabetis seperti titik, koma, tanda tanya, tanda seru, serta simbol-simbol seperti "@" untuk nama pengguna, hashtag, emoticon, dan alamat web dihapus.
3. Folding Cases: Mengubah karakter alfabet yang telah dibersihkan menjadi huruf kecil, yang dikenal sebagai folding case.
4. Tokenizing: Memecah kalimat menjadi kata-kata yang membentuknya, disebut term atau token, berdasarkan spasi.
5. Filtering: Menghapus kata-kata yang tidak penting atau tidak bermakna untuk analisis sentimen, seperti yang, dengan, di, ke, dan tetapi.
6. Stemming: Mengubah kata berimbuhan menjadi kata dasar.

### 2.6 Support Vector Machine

Menurut studi yang dilakukan oleh Suryati dan rekan pada tahun 2023, Support Vector Machine (SVM) diakui sebagai salah satu algoritma machine learning yang sangat terkenal dalam hal klasifikasi. Dalam dekade terakhir, SVM telah menjadi metode yang sangat efektif untuk mengidentifikasi pola klasifikasi dan telah meraih tingkat keberhasilan yang signifikan di berbagai bidang. Oleh karena itu, minat dari komunitas machine learning terus meningkat dalam mempelajari dan mengembangkan SVM karena reputasinya yang baik dalam menyelesaikan berbagai tantangan pembelajaran. SVM berfungsi dengan mencari hyperplane optimal yang mampu memisahkan dua kelas dalam ruang input. Algoritma klasifikasi SVM menggunakan data latihannya untuk membentuk model klasifikasi, yang kemudian dapat digunakan untuk memprediksi kelas data baru, yang dikenal sebagai data pengujian. .

## 3. METODE PENELITIAN

### 3.1 Pengumpulan Data

Metode pengumpulan data yang diterapkan dalam penelitian "Analisis Sentimen Tingkat Kepuasan Pelanggan Terhadap Layanan Kurir J&T Express Menggunakan Algoritma Support Vector Machine (SVM) Berdasarkan Review Pengguna Di Google Playstore" mencakup:

1. Studi Literatur: Peneliti mengumpulkan Data dari jurnal atau karya tulis ilmiah yang relevan, sehingga dapat membantu peneliti dalam menambah referensi sesuai dengan topik yang diangkat.
2. Filter Data: Peneliti menggunakan filtering Data dari Dataset yang akan digunakan, untuk mendapatkan variabel dan Data yang dapat digunakan.

### 3.2 Metode Implementasi

Penerapan model yang dijelaskan oleh peneliti menggunakan metode CRISP-DM (Cross Industry Standard Process for Data Mining) melibatkan beberapa tahap, sebagai berikut:

1. Pemahaman Bisnis (Business Understanding): Pada tahap ini, penelitian memahami dan mengumpulkan data melalui metode Text Mining di Google Play Store.
2. Pemahaman Data (Data Understanding): Data dianalisis untuk memastikan apakah sudah memadai dan layak untuk diproses atau apakah perlu dilakukan pengumpulan data tambahan.
3. Persiapan Data (Data Preparation): Melalui proses Text Preprocessing, data dibersihkan dan diolah sehingga menghasilkan data yang berkualitas. Selanjutnya, dilakukan labeling atau pelabelan kata-kata sebagai positif dan negatif.

4. **Modeling:** Pada tahap ini, dilakukan pembobotan kata positif dan negatif menggunakan TF-IDF. Algoritma Support Vector Machine kemudian diterapkan pada data latih dan data uji untuk mendapatkan tingkat akurasi.
5. **Evaluasi (Evaluation):** Setelah akurasi diperoleh, algoritma Support Vector Machine yang telah dilatih dan diuji dievaluasi menggunakan Confusion Matrix untuk menilai sejauh mana kemampuannya dalam mengklasifikasi data. Hasil model yang sudah dilatih dan diuji dapat disimpan menggunakan Python pickle.
6. **Implementasi (Deployment):** Selanjutnya, untuk dapat diakses oleh pengguna, Analisis Sentimen akan di-deploy atau di-hosting menggunakan web framework Streamlit, dan Deta Space sebagai Cloud Database

#### 4. PEMBAHASAN

##### 4.1 Business Understanding

Tujuan penelitian ini adalah untuk menganalisis ulasan pengguna terkait aplikasi J&T Express dengan tujuan menilai apakah umumnya ulasan tersebut bersifat positif atau negatif. Hal ini dilakukan untuk membantu J&T Express mengidentifikasi masalah yang mungkin timbul dalam layanannya. Dengan pemahaman atas masalah-masalah tersebut, diharapkan J&T Express dapat meningkatkan mutu layanannya sehingga memberikan pengalaman yang lebih baik bagi pengguna. Rating rendah untuk aplikasi J&T Express (2,4 dari 5 bintang) menandakan adanya keluhan atau ketidakpuasan dari pengguna, sehingga sangat penting bagi J&T Express untuk mengenali dan memahami masalah yang muncul.

Peneliti mengumpulkan dataset ulasan pengguna J&T Express dan melakukan analisis menggunakan metode Support Vector Machine untuk mengklasifikasikan apakah ulasan tersebut memiliki sentimen positif atau negatif. Selain itu, penelitian ini juga bertujuan untuk mendapatkan wawasan baru dari dataset ulasan pengguna J&T Express dan mengevaluasi performa klasifikasi yang dihasilkan menggunakan metode Support Vector Machine.

##### 4.2 Data Understanding

Data dihimpun melalui teknik Web Scraping dari ulasan pengguna aplikasi J&T Express di Google Play Store, menggunakan perpustakaan Python yang disebut `google_play_scraper`. Awalnya, 9000 data diambil karena dianggap paling relevan. Namun, dilakukan penyaringan data dengan menghapus kata "aplikasi" atau "apk" agar fokus pada layanan kurir, sehingga jumlah data berkurang menjadi 7815.

##### 4.3 Data Preparation

Sebelum data disertakan dalam proses pemodelan, dilakukan tahap pembersihan atau penyempurnaan data. Langkah ini bertujuan untuk menciptakan model dan akurasi optimal sambil mempermudah manajemen data oleh peneliti. Dalam konteks ini, peneliti memilih empat sampel data untuk dipersiapkan. Salah satu dari empat data yang terdapat di dalam bagian "isi" dapat dilihat pada Gambar 1..

content
Pelayanan sangat buruk, barang telah tertunda selama 2 hari, membuat saya merasa kecewa.
Sangat mengecewakan, setelah menunggu lama, paket justru hilang.
Layanan nya murah dan dijamin cepat
saya kecewa dengan pelayanan jnt, buruk!!!

Gambar 1 data sampel pada content

Berikut tahapan pada Data Preparation yang telah peneliti implementasi:

##### 4.3.1 Case Folding

Untuk membuat pengelolaan data lebih mudah, proses Case Folding mengubah teks menjadi huruf kecil dan menghapus karakter tertentu. Sebagai contoh, kata "pelayanan" akan dikecilkan menjadi "pelayanan".

caseFolding
pelayanan sangat buruk, barang telah tertunda selama 2 hari, membuat saya merasa kecewa.
sangat mengecewakan, setelah menunggu lama, paket justru hilang.
layanan nya murah dan dijamin cepat
saya kecewa dengan pelayanan jnt, buruk!!!

Gambar 2 Proses Case Folding

4.3.2 *Cleansing*

Pada tahap Cleansing dilakukan dengan membersihkan teks dari noise atau elemen yang tidak diperlukan, meliputi angka, tanda baca, emoji dan spasi berlebih. Sebagai contoh kata “2 hari” angkanya di hilangkan menjadi “hari”, dan “jelekkk!!!” tanda serunya dihilangkan menjadi “jelekkk”

cleansing
pelayanan sangat buruk barang telah tertunda selama hari membuat saya merasa kecewa
sangat mengecewakan setelah menunggu lama paket justru hilang
layanannya murah dan dijamin cepat
saya kecewa dengan pelayanan jnt buruk

Gambar 3 Proses Cleansing

4.3.3 *Stemming*

Pada tahap stemming dilakukan dengan menghilangkan imbuhan atau akhiran kata dalam teks sehingga hanya menyisakan bentuk dasarnya. Sebagai contoh kata “pelayanan” imbuhnya di hilangkan menjadi “layan”.

stemmer
layan sangat buruk barang telah tunda lama hari buat saya rasa kecewa
sangat kecewa telah tunggu lama paket justru hilang
layan nya murah dan jamin cepat
saya kecewa dengan layan jnt buruk

Gambar 4 Proses Stemming

4.3.4 *Slang Word Standardization*

Pada tahap slang word standardization dilakukan dengan mengubah kata gaul “slang” menjadi kata formal menggunakan kamus slang word yang sudah peneliti buat. Sebagai contoh kata “jelekkk” di ubah kata gaulnya menjadi formal “jelek”

slang
layan sangat buruk barang telah tunda lama hari buat saya rasa kecewa
sangat kecewa telah tunggu lama paket justru hilang
layan nya murah dan jamin cepat
saya kecewa dengan layan jnt buruk

Gambar 5 Proses Slang Word Standardization

4.3.5 *Stop Word Removal*

Pada tahap stop word removal dilakukan dengan menghapus kata-kata yang umum dan tidak memiliki informasi penting dalam teks. Sebagai contoh kata “saya” di hilangkan.

stopword
layan buruk barang tunda kecewa
kecewa tunggu paket hilang
layan murah jamin cepat
kecewa layan jnt buruk

Gambar 6 Proses Stop Word Removal

4.3.6 *Unwanted Word Removal*

Pada tahap ini, kata-kata yang telah dibuat oleh peneliti dibuang, termasuk nama bulan seperti "januari" dan "jan" seterusnya.

stopword
layan buruk barang tunda kecewa
kecewa tunggu paket hilang
layan murah jamin cepat
kecewa layan jnt buruk

Gambar 7 Proses Unwanted Word Removal

#### 4.3.7 Short Word Removal

Pada tahap pertama, kata-kata yang di bawah tiga huruf dihilangkan karena tidak memiliki arti yang signifikan.

#### 4.3.8 Tokenizing

Pada tahap tokenizing dilakukan dengan membuat Split Text (memisahkan kata-kata) menjadi per kata untuk masing-masing ulasan.

#### 4.3.9 Labeling Data

Pada Tahap labeling Data dilakukan pelabelan Data agar mengetahui kalimat bersifat positif atau negatif dengan menggunakan kamus positif dan negatif yang sudah dibuat oleh penulis lalu dihitung jumlahnya, jika suatu kalimat mengandung kata positif akan +1 dan jika mengandung kata negatif akan -1.

#### 4.3.10 Term Frequency-Inverse Document Frequency (TF-IDF)

Tahap TF-IDF (Term Frequency-Inverse Document Frequency) dilakukan dengan menghitung bobot nilai setiap kata dalam teks. Bobot ini menunjukkan seberapa sering kata tersebut muncul dalam dokumen tertentu (Term Frequency) dan seberapa umum atau jarang kata tersebut muncul di setiap dokumen.

#### 4.3.11 Modeling

Berikut adalah tahapan dari modeling menggunakan library dan perhitungan manual.

#### 4.3.12 Split Data

Pada tahap modeling menggunakan library, pertama-tama akan dilakukan split Data (pemisahan Data) menjadi Data uji dan Data test pembagian Data ini menjadi 90% untuk Data uji dan 10% untuk Data test.

#### 4.3.13 Penerapan Library Support Vector Machine

Terakhir akan dilakukan tahap pemodelan dengan Support Vector Machine kernel liner untuk melihat akurasi yang di hasilkan oleh Support Vector Machine.

#### 4.3.14 Perhitungan Manual

Dimulai dengan melakukan perhitungan secara manual, data akan dipisahkan menjadi empat sampel, terdiri dari tiga data untuk pelatihan dan satu data untuk pengujian. Setelah itu, tf-idf akan diterapkan pada data pelatihan, seperti yang terlihat dalam ilustrasi berikut. Selanjutnya, TF (Frekuensi Istilah) akan dihitung dan dijumlahkan untuk setiap dokumen, sebagaimana tampak dalam gambar berikutnya. Tahapan berikutnya melibatkan normalisasi tf menggunakan rumus tertentu. Kemudian, kami akan menghitung TF (Term Frequency) dan menjumlahkannya untuk setiap dokumen, sesuai yang terlihat dalam gambar berikutnya. Langkah selanjutnya melibatkan normalisasi tf menggunakan rumus tertentu. Proses berikutnya adalah mencari df dengan mengidentifikasi dokumen yang memuat suatu kata. Sebagai contoh, untuk kata "layan", kata tersebut muncul di dalam dua dokumen, sehingga hasil df-nya adalah 2. Langkah terakhir dalam perhitungan adalah menghitung TF-IDF dengan mengalikan tf dengan hasil idf, seperti yang ditunjukkan dalam rumus  $tf \ idf = tf \times idf$ . Berikut adalah penyelesaian untuk kata "layan" berdasarkan rumus tf-idf yang telah disebutkan sebelumnya, yaitu  $tf \ idf = 0,2 \times 0,301 = 0,06$ .

Setelah tahap tf-idf, langkah selanjutnya adalah melakukan pemodelan menggunakan Support Vector Machine dengan kernel linear. Proses ini membutuhkan nilai support vector (a) dan bobot vektor (w) untuk menentukan hyperplane (b). Kemudian, dilakukan proses pelatihan pemodelan data menggunakan data latih yang telah diberi bobot sebelumnya. Hasil pembobotan kata pada data latih diubah menjadi bentuk vektor, dengan dokumen yang

berlabel positif direpresentasikan dengan angka 1, dan dokumen yang berlabel negatif direpresentasikan dengan angka -1. Berikut adalah representasi vektor pada dokumen D1, D2, dan D3.

#### 4.4 Evaluation

Dalam tahap evaluasi, model diuji untuk kinerjanya menggunakan dua proses, yaitu laporan klasifikasi dan matriks kebingungan. Laporan klasifikasi memberikan gambaran tentang bagaimana model melakukan prediksi sentimen dari teks, dengan menggunakan metrik evaluasi seperti presisi, recall, dan f1-score untuk mengevaluasi kualitas prediksi sentimen. Presisi mengukur akurasi model dalam memprediksi sentimen positif atau negatif, sementara recall menunjukkan seberapa baik model dalam menemukan dan mengklasifikasikan dengan benar sentimen positif atau negatif. F1-score merupakan skor yang menggabungkan presisi dan recall untuk memberikan penilaian keseluruhan tentang kinerja model. Selain itu, Classification Report juga memberikan informasi tentang support, yang mengindikasikan jumlah sampel yang termasuk dalam setiap kategori sentimen.

Di samping itu, terdapat matriks kebingungan (confusion matrix), yang digunakan untuk mengukur kualitas prediksi model antara sentimen positif dan negatif. Matriks ini membandingkan hasil prediksi model dengan label sebenarnya dari data yang dianalisis. Matriks kebingungan terdiri dari empat bagian: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). True Positive menunjukkan jumlah prediksi yang benar bahwa sentimen adalah positif, sementara True Negative menunjukkan jumlah prediksi yang benar bahwa sentimen adalah negatif. False Positive adalah jumlah prediksi yang salah bahwa sentimen adalah positif, sedangkan False Negative menunjukkan jumlah prediksi yang salah bahwa sentimen adalah negatif. Jika jumlah TP dan TN lebih besar daripada FP dan FN, itu menunjukkan bahwa model mampu mengenali sentimen positif dan negatif dengan baik..

#### 4.5 Deployment

Tahap Deployment merupakan langkah akhir dalam metodologi CRISP-DM, di mana proses Data mining diimplementasikan ke dalam sebuah situs web. Peneliti menggunakan Streamlit sebagai *platform web hosting* dan *Data Space* sebagai penyimpanan cloud untuk data.

## 5. PENUTUP

### 5.1 Kesimpulan

Hasil penelitian dijelaskan secara detail dalam bagian diskusi sebagai berikut:

1. Berdasarkan kinerja metode Support Vector Machine (SVM), SVM berhasil mencapai tingkat akurasi sebesar 96,56%. Angka tersebut telah dibulatkan menjadi 97% untuk memudahkan pemahaman. Menurut para peneliti, kinerja ini dianggap memuaskan.
2. Dari analisis sentimen terhadap aplikasi J&T Express, ditemukan lebih banyak polaritas negatif daripada positif. Sentimen negatif yang terakumulasi mengindikasikan keluhan terkait "pelayanan buruk dan kecewa karena pengiriman paket yang lambat, kesalahan pengiriman paket, dan kehilangan paket." Di sisi lain, sentimen positif yang terkumpul menunjukkan pujian terhadap "pelayanan yang baik, kurir yang efisien karena pengiriman yang cepat, dan barang yang sesuai."

### 5.2 Saran

1. Untuk penelitian mendatang, disarankan untuk mempertimbangkan penggunaan dataset yang lebih besar guna mendapatkan informasi yang lebih komprehensif dan melibatkan variasi yang lebih luas. Selain itu, direkomendasikan untuk menggunakan kamus yang terkini agar data tetap relevan, dan mengevaluasi penggunaan algoritma alternatif sebagai pembanding.
2. Bagi J&T Express, disarankan untuk memperhatikan hasil analisis dan implementasi dari penelitian ini. Diharapkan perusahaan dapat memperbaiki layanannya berdasarkan sentimen yang terungkap dari ulasan pengguna, sehingga J&T Express dapat mencapai kemajuan yang lebih baik di masa mendatang.

## DAFTAR PUSTAKA

- [1] Aziz, A. (2022). Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength. Dalam Jurnal Sains Komputer & Informatika (J-SAKTI (Vol. 6, Nomor 1).
- [2] Deta Base – Space Docs (2023), Diakses pada 19 Juni 2023, dari <https://deta.space/docs/en/build/reference/deta-base>
- [3] Getting Started – Visual Studio Code Docs (2023), diakses pada 19 juni 2023, dari <https://code.visualstudio.com/docs>
- [4] Gifari, O. I., Adha, M., Rifky Hendrawan, I., Freddy, F., & Durrand, S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan *Support Vector Machine*. Jifotech (Journal Of Information Technology, 2(1).
- [5] Hasna, S. K. (2021). Analisis Sentimen *Data* Ulasan Menggunakan Algoritma *Support Vector Machine*. Tugas Akhir, Universitas Islam Indonesia
- [6] Kurniawan, D., & Yasir, D. M. (2022). Optimization Sentiment Analysis Using Crisp-Dm And Naïve Bayes Methods Implemented On Social Media. 6, 74–84.
- [7] Locarso, G. K. (2022). Analisis Sentimen Review Aplikasi Pedulilindungi Pada *Google Play Store* Menggunakan Nbc. Jurnal Teknik Informatika Kaputama (JTIK), 6(2).
- [8] Pasek, P., Mahawardana, O., Sasmita, G. A., Agus, P., & Pratama, E. (2022). Analisis Sentimen Berdasarkan Opini dari Media Sosial Twitter terhadap “Figure Pemimpin” Menggunakan Python. Dalam JITTER-Jurnal Ilmiah Teknologi dan Komputer (Vol. 3, Nomor 1).
- [9] Pratama, A. E., Ariesta, A., & Gata, G. (2022). Analisis Sentimen Masyarakat terhadap Tim Nasional Indonesia pada Piala AFF 2020 Menggunakan Algoritma K-Nearest Neighbors The researcher uses the Cross-Industry Standard Process for *Data Mining* (CRISP-DM) method and implements the K-Nearest. Jurnal TICOM: Technology of Information and Communication, 10(3), 187–196.
- [10] Ramadhani, S. H., & Wahyudin, M. I. (2022). Analisis Sentimen Terhadap Vaksinasi Astra Zeneca pada Twitter Menggunakan Metode Naïve Bayes dan K-NN. Jurnal Teknologi Informasi dan Komunikasi, 6(4), 2022.
- [11] Suryati, E., Ari Aldino, A., Penulis Korespondensi, N., & Suryati Submitted, E. (2023). Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma *Support Vector Machine* (SVM). 4(1), 96–106.
- [12] *Streamlit* Documentation – *Streamlit* Docs (2023), Diakses pada 19 Juni 2023, dari <https://docs.streamlit.io/>