**ANALISIS SENTIMEN ULASAN PRODUK APLIKASI E-COMMERCE LAYANAN TRANSPORTASI INDONESIA**

TUGAS BESAR DATA MINING

Oleh

Revan Yuda Febriansyah 714220070

Bargana Kukuh Raditya 714220013

Ruth Diana Purnamasari Sagala 714220042

Gilang Andhika Buwana 714220046



**DIPLOMA IV TEKNIK INFORMATIKA**

**SEKOLAH VOKASI**

**UNIVERSITAS LOGISTIK DAN BISNIS INTERNASIONAL**

**BANDUNG**

**2025**

# HALAMAN PERNYATAAN ORISINALITAS

Tugas besar ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku

Yang menyatakan,

Nama : …………..

NIM : …………..

Tanda Tangan : …………..

Tanggal : …………..

Mengetahui

Ketua :................................... (.......tanda tangan. )

Pembimbing I :................................... (.......tanda tangan. )

# KATA PENGANTAR

Pedoman penulisan skripsi sebagai hasil dari Tugas Penelitian Data atau Proyek Akhir pada Program Studi Informatika dibuat untuk membantu mahasiswa yang sedang menyusun laporan tugas akhir, baik itu berupa laporan kemajuan maupun laporan akhir penelitian. Skripsi Tugas Akhir Program Studi Informatika ini merupakan karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Informatika dari Universitas Logistik Bisnis Internasional.

Karya ini akan menjadi bagian dari koleksi Perpustakaan S Universitas Logistik Bisnis Internasional Bandung sebagai suatu karya ilmiah yang dihasilkan oleh sivitas akademika ULBI. Berdasarkan keperluan tersebut, maka keseragaman format dan penggunaan tata bahasa Indonesia yang baik dan benar merupakan suatu keharusan dalam laporan akhir tugas besar tersebut.

Oleh karena itu, dalam pedoman ini diuraikan berbagai hal yang berkaitan dengan struktur karya ilmiah dan teknik penulisannya. Pedoman ini disusun sebagai hasil adaptasi dari berbagai sumber pedoman penulisan tugas akhir dari berbagai universitas, yang kemudian disesuaikan dengan kebutuhan Program Studi Informatika. Dengan demikian, akan terdapat kesamaan dengan pedoman karya tulis ilmiah lain baik dari dalam negeri maupun mancanegara. Beberapa penyederhanaan dan modifikasi juga diberikan demi mempertimbangkan substansi dan kemudahan dalam penulisan.

Akhir kata, penulis dengan segala kerendahan hati bersedia menerima kritik dan masukan yang membangun demi penyempurnaan pedoman penulisan Laporan Akhir Program Studi Informatika ini.

Bandung, Juli 2025

# HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Logistik Bisnis Internasional, saya yang bertanda tangan di bawah ini:

Nama : ...............................................

NIM : ...............................................

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Logistik Bisnis Internasional, Hak Bebas Royalti Noneksklusif (*Non- exclusive Royalti Free Right*) atas karya ilmiah saya yang berjudul:

......................................................................................................................................................

......................................................................................................................................................

Beserta perangkat yang ada (jika diperlukan). Dengan Hak ini Universitas Logistik Bisnis Internasional Hayati berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya. Dibuat di : ..............................

Pada tanggal : ..............................

Yang menyatakan

( )

# DAFTAR ISI

[HALAMAN PERNYATAAN ORISINALITAS 2](#_Toc204334788)

[KATA PENGANTAR 3](#_Toc204334789)

[HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS 4](#_Toc204334790)

[DAFTAR ISI 5](#_Toc204334791)

[BAB 1 PENDAHULUAN 7](#_Toc204334792)

[**1.1** **Latar Belakang** 7](#_Toc204334793)

[**1.2** **Rumusan Masalah** 7](#_Toc204334794)

[**1.3** **Tujuan Penelitian** 8](#_Toc204334795)

[**1.4** **Manfaat Penelitian** 8](#_Toc204334796)

[**1.5** **Ruang Lingkup** 8](#_Toc204334797)

[BAB II TINJAUAN PUSTAKA 10](#_Toc204334798)

[**2.1** **Kajian Teori** 10](#_Toc204334799)

[2.1.1 Data Mining 10](#_Toc204334800)

[2.1.2 Machine Learning 10](#_Toc204334801)

[2.1.3 Naive Bayes 10](#_Toc204334802)

[2.1.4 Analisis Sentimen 10](#_Toc204334803)

[2.1.5 Ulasan Pengguna 10](#_Toc204334804)

[2.1.6 Transportasi Online 10](#_Toc204334805)

[2.1.7 Klasifikasi Teks 10](#_Toc204334806)

[**2.2** **State Of The Art** 11](#_Toc204334807)

[BAB III METODOLOGI PENELITIAN 12](#_Toc204334808)

[**3.1** **Tahapan Penelitian** 12](#_Toc204334809)

[**3.2** **Deskripsi Dataset** 12](#_Toc204334810)

[**3.3** **Algoritma** 12](#_Toc204334811)

[**3.4** **Evaluasi Kinerja** 13](#_Toc204334812)

[BAB IV HASIL DAN PEMBAHASAN 14](#_Toc204334813)

[**4.1** **Visualisasi dan EDA** 14](#_Toc204334814)

[4.1.1 Distribusi Label Sentimen 14](#_Toc204334815)

[4.1.2 Word Cloud 14](#_Toc204334816)

[4.1.3 Statistik Deskriptif 15](#_Toc204334817)

[**4.2** **Hasil *Preprocessing* dan *Feature Engineering*** 15](#_Toc204334818)

[**4.3** **Hasil *Preprocessing* dan Pemodelan** 16](#_Toc204334819)

[4.3.1 Confusion Matrix 17](#_Toc204334820)

[**4.4** **Interpretasi Hasil dan Analisis Komparatif** 19](#_Toc204334821)

[**4.5** **Analisis Keunggulan dan Keterbatasan** 20](#_Toc204334822)

[BAB V KESIMPULAN DAN SARAN 21](#_Toc204334823)

[**5.1** **Kesimpulan** 21](#_Toc204334824)

[**5.2** **Jawaban atas Rumusan Masalah** 21](#_Toc204334825)

[**5.3** **Saran** 21](#_Toc204334826)

[DAFTAR PUSTAKA 23](#_Toc204334827)

# BAB 1 PENDAHULUAN

## **Latar Belakang**

Pada era digital saat ini, perkembangan teknologi informasi yang pesat telah mendorong digitalisasi di berbagai sektor, termasuk sektor transportasi berbasis teknologi modern [1]. Di Indonesia, ada banyak aplikasi *e-commerce* yang menyediakan berbagai layanan untuk meningkatkan efisiensi dan efektivitas. Salah satunya Gojek yang telah menjadi bagian penting dalam kehidupan masyarakat yang memiliki berbagai layanan dari mulai pemesanan transportasi, pengiriman barang, pemesanan makanan, hingga pembayaran digital. Hal ini membuat industri transportasi darat lebih terbuka dan memiliki daya saing tinggi dalam mendapatkan konsumen [2].

Seiring meningkatnya jumlah pengguna aplikasi tersebut, *review* atau ulasan yang diberikan oleh pengguna juga semakin banyak [3]. Ulasan tersebut dapat berisi berbagai macam tanggapan, baik positif negatif, maupun netral yang mencerminkan kepuasan atau kekecewaan pengguna terhadap layanan yang mereka terima. Ulasan tersebut menjadi data yang sangat berharga, bukan bagi perusahaan untuk meningkatkan kualitas layanan nya, tetapi juga bagi calon pengguna sebagai bahan pertimbangan sebelum menggunakan layanan tersebut [3], [4].

Dalam beberapa tahun terakhir, media sosial menjadi sumber utama bagi para peneliti untuk menggali opini terkait berbagai isu yang dihadapi oleh masyarakat. Khususnya, platform seperti X telah berkembang menjadi ruang terbuka bagi pengguna untuk menyampaikan pendapat, keluhan, maupun apresiasi terhadap suatu layanan, termasuk layanan transportasi berbasis aplikasi [5].

Besarnya jumlah ulasan dengan jutaan tweet yang dipublikasikan setiap harinya, menjadi tidak mungkin untuk menganalisisnya secara manual. Oleh karna itu, dibutuhkan metode otomatis untuk megolah dan memahami opini publik dari banyaknya ulasan tersebut [5]. Analisis sentimen merupakan proses pengolahan bahasa alami yang bertujuan untuk mengidentifikasi opini atau perasaan pengguna terhadap suatu topik, terutama terhadap layanan transportasi dari aplikasi *e-commerce*. Penelitian ini diharapkan dapat memberikan gambaran umum mengenai presepsi masyarakat, serta menjadi dasar dalam pengambilan keputusan yang lebih berbasis data.

## **Rumusan Masalah**

1. Bagaimana tingkat akurasi algoritma Naive Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi *e-commerce* layanan transportasi di Indonesia berdasarkan data dari media sosial X (Twitter)?
2. Metode representasi teks manakah yang memberikan kinerja terbaik ketika dikombinasikan dengan algoritma Naive Bayes untuk analisis sentimen ulasan aplikasi transportasi online, apakah TF-IDF, Bag-of-Words (BoW), Word2Vec, atau TF-IDF dengan seleksi fitur Chi-Squared?
3. Bagaimana distribusi sentimen (positif, negatif, dan netral) dari ulasan pengguna aplikasi Gojek di media sosial X dan aspek apa saja yang paling mempengaruhi sentimen tersebut?

## **Tujuan Penelitian**

1. Menganalisis dan mengevaluasi kinerja algoritma Naive Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi *e-commerce* layanan transportasi berdasarkan data ulasan dari media sosial X (Twitter).
2. Membandingkan efektivitas berbagai metode representasi teks (TF-IDF, Bag-of-Words, Word2Vec, dan TF-IDF + Chi-Squared) dalam meningkatkan akurasi klasifikasi sentimen menggunakan algoritma Naive Bayes.
3. Mengidentifikasi dan menganalisis pola sentimen masyarakat Indonesia terhadap layanan transportasi online Gojek berdasarkan ulasan di media sosial X untuk memberikan insight mengenai persepsi publik terhadap layanan tersebut.

## **Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat baik dari segi akademis maupun praktis. Adapun manfaat yang dapat diperoleh antara lain:

1. Memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen, text mining, dan machine learning.
2. Menjadi referensi bagi mahasiswa maupun peneliti lain yang tertarik untuk melakukan kajian serupa mengenai pemrosesan bahasa alami terhadap data media sosial berbahasa Indonesia.
3. Memberikan gambaran umum tentang persepsi masyarakat terhadap layanan transportasi berbasis aplikasi di Indonesia, berdasarkan analisis ulasan pengguna di media sosial X.
4. Membantu perusahaan penyedia layanan transportasi dalam mengevaluasi kualitas layanan, mendeteksi masalah umum yang dikeluhkan pelanggan, serta Menyusun strategi peningkatan layanan berdasarkan opini publik.

## **Ruang Lingkup**

Ruang lingkup dari penelitian ini dibatasi pada analisis sentimen terhadap ulasan pengguna aplikasi *e-commerce* layanan transportasi yang berasal dari platform X. Penelitian difokuskan pada beberapa hal berikut:

1. **Sumber Data**

Data ulasan yang dianalisis berasal dari media sosial X tweet publik yang memuat opini pengguna mengenai aplikasi *e-commerce* layanan transportasi di Indonesia yaitu Gojek. Data disambil dalam periode tertentu yang telah ditentukan.

1. **Bahasa Ulasan**

Hanya tweet yang menggunakan Bahasa Indonesia yang dianalisis, mengingat fokus penelitian berada pada konteks lokal pengguna di Indonesia.

1. **Metode Analisis**

Analisis sentimen dilakukan dengan menggunakan algoritma klasifikasi Naive Bayes yang dikombinasikan dengan beberapa pendekatan representasi teks, yaitu TF-IDF, Bag-of-Words (BoW), Word2Vec, serta TF-IDF dengan seleksi fitur menggunakan metode Chi-Squared. Pendekatan ini digunakan untuk mengidentifikasi dan mengelompokkan sentimen dalam ulasan menjadi kategori positif, negatif, atau netral.

1. **Jenis Sentimen**

Penelitian ini mengklasifikasikan ulasan ke dalam tiga kategori utama sentimen, yaitu positif, negatif, dan netral, sesuai dengan konteks dan isi tweet.

1. **Keterbatasan Studi**

Penelitian ini mencakup analisis mendalam terhadap faktor-faktor penyebab munculnya sentimen, tidak melakukan analisis per individu pengguna, dan tidak mencakup platform media sosial lain selain X.

# BAB II TINJAUAN PUSTAKA

## **Kajian Teori**

### Data Mining

Data mining merupakan proses analisis terhadap kumpulan data berukuran besar untuk menemukan pola-pola tersembunyi yang bernilai guna. Aktivitas data mining dapat mencakup beberapa tahapan penting seperti klasifikasi, integrasi, transformasi, diskritisasi, dan evaluasi pola data. Tekni ini banyak dimanfaatkan dalam berbagai bidang, seperti pemasaran, deteksi kecurangan, penjualan, analisis sentimen hingga pengembangan suatu produk [6].

### Machine Learning

Machine Learning adalah cabang dari kecerdasan buatan (Artificial Intelligence) yang berfokus pada pengembangan algoritma dan teknik yang memungkinkan komputer belajar dari data dan melakukan prediksi atau pengambilan keputusan tanpa deprogram secara eksplisit [7].

### Naive Bayes

Naive Bayes adalah metode klasifikasi yang sering digunakan dalam *machine learning*, khususnya untuk klasifikasi teks seperti analisis sentimen, deteksi spam, dan pengelompokan dokumen. Metode ini didasarkan pada prinsip Teorema Bayes dengan asumsi sederhana bahwa setiaip fitur dalam data bersifat independen sehingga disebut “naive” atau sederhana [7].

### Analisis Sentimen

Analisis Sentimen merupakan proses untuk mengidentifikasi, mengekstrak, dan mengklasifikasikan emosi atau opini seseorang terhadap suatu objek (seperti produk, kebijakan, layanan, ataupun peristiwa) dari sebuah teks apakah bersifat positif, negatif, atau netral [5].

### Ulasan Pengguna

Ulasan pengguna merupakan bentuk umpan balik yang diberikan oleh pengguna setelah Menggunakan suatu produk, layanan, atau aplikasi, biasanya berbentuk teks atau penilaian numerik (rating) [4]. Ulasan pengguna menjadi sumber informasi penting yang dapat memengaruhi keputusan konsumen lain serta menjadi masukin berharga bagi penyedia layanan untuk meningkatkan kualitas produk [2].

### Transportasi Online

Transportasi online adalah layanan transportasi berbasis teknologi digital, di mana pengguna dapat memesan kendaraan secara langsung melalui aplikasi yang terhubung ke internet [2]. Transportasi online menawarkan kemudahan, kecepatan, dan efisiensi dalam sistem pemesanan kendaraan. Kehadiran layanan ini telah mengubah pola mobilitas masyarakan modern, karena memberikan akses transportasi yang fleksibel dan ekonomis contohnya layanan gojek.

### Klasifikasi Teks

Klasifikasi teks adalah bagian dari teknik klasifikasi dalam data mining, di mana proses utamanya adalah memetakan sebuah data teks ke dalam kelas-kelas tertentu yang telah ditentukan sebelumnya [6]. Tujuan dari klasifikasi teks adalah untuk mengelompokkan dokumen atau kalimat ke dalam kategori seperti positif, negatif, atau netral dalam analisis sentimen.

## **State Of The Art**

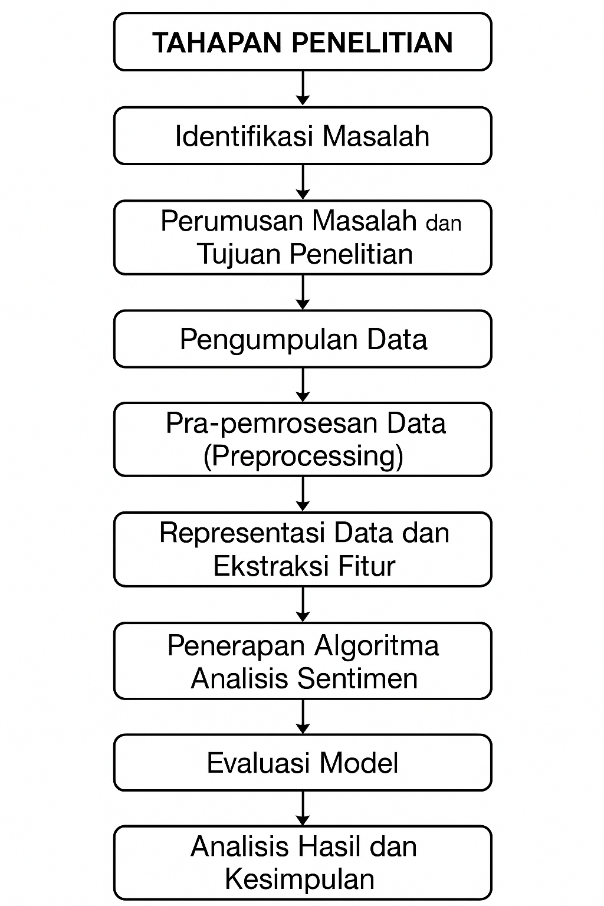
Dalam beberapa tahun terakhir, penelitian analisis sentimen berbasis text mining menunjukkan evolusi signifikan dari model klasik yang ringan komputasi seperti Naïve Bayes, SVM, dan K-Nearest Neighbor ke model yang memanfaatkan representasi vektor kaya dan arsitektur pembelajaran mendalam, termasuk CNN, LSTM, model hibrida CNN-LSTM, serta keluarga model berbasis Transformer (BERT, IndoBERT, IndoBERTweet) yang dilatih atau domain-adapted pada data media sosial [8], [9], [10].

Di ranah layanan digital dan transportasi online, studi awal banyak mengandalkan pendekatan sederhana (Naive Bayes atau API analitik siap-pakai) untuk menilai persepsi pengguna terhadap platform seperti GOJEK; penelitian-penelitian ini umumnya fokus pada klasifikasi polaritas tunggal (positif/negatif) atau ringkasannya saja. Tren terbaru bergerak ke analisis yang lebih kaya: multi-label / multi-aspect sentiment analysis untuk memetakan dimensi layanan (akses, waktu, kenyamanan, keamanan, customer service, dsb.) sehingga temuan dapat langsung ditindaklanjuti secara operasional oleh penyedia layanan [8], [11], [12].

Mencermati lanskap tersebut, masih terdapat research gap pada kajian berbahasa Indonesia yang: (1) secara langsung membandingkan beberapa representasi teks klasik (TF-IDF, BoW), embedding distribusional (Word2Vec), dan feature selection (Chi-Squared) dalam satu kerangka yang konsisten; (2) menguji kinerja model Naive Bayes pada dataset ulasan media sosial berukuran kecil yang lazim ditemui peneliti pemula; dan (3) mengeksplor strategi penanganan ketidakseimbangan kelas sebelum/selama pemodelan. Penelitian Anda diarahkan untuk mengisi celah ini melalui eksperimen terkontrol pada data tweet Bahasa Indonesia tentang layanan GOJEK yang dikumpulkan periode 21 Mei–02 Juli 2025, dengan membandingkan empat skema representasi (TF-IDF, BoW, Word2Vec, TF-IDF+Chi-Squared) pada algoritma Naive Bayes dan metode interpretabilitas di masa depan.

# BAB III METODOLOGI PENELITIAN

## **Tahapan Penelitian**



## **Deskripsi Dataset**

Dataset yang digunakan dalam penelitian ini diperoleh dari media sosial X. data disambil Menggunakan teknik web scraping melalui X API (menggunakan library seperti snscrape) dengan kata kunci tertentu seperti “Gojek” dan “aplikasi gojek”. Pengambilan data dilakukan pada rentang waktu 21 Mei 2025 hingga 02 Juli 2025. Total tweet yang berhasil dikumpulkan berjumlah sebanyak 282 ulasan. Dataset disimpan dalam format CSV dan terdiri dari beberapa kolom.

## **Algoritma**

Penelitian ini menggunakan pendekatan klasifikasi teks untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi Gojek yang dikumpulkan dari media sosial X. Algoritma utama yang digunakan adalah Naive Bayes, yang dikenal sebagai metode klasifikasi berbasis probabilistik sederhana namun efektif untuk teks.

Sebelum diterapkan ke model, data ulasan terlebih dahulu diproses melalui beberapa tahapan text preprocessing yaitu dimuali dari Case Folding, Cleanshing, Tokenizing, Stopword Removal, dan Stemming. Setelah preprocessing, data direpresentasikan dalam bentuk numerik menggunakan 4 pendekatan representasi teks, yang kemudian digunakan untuk melatih model Naive Bayes:

1. **Naive Bayes + TF-IDF (Term Frequency-Inverse Document Frequency)**

TF-IDF digunakan untuk merepresentasikan teks dengan memperhatikan seberapa penting sebuah kata dalam dokumen. Pendekatan ini memperkuat kata-kata yang spesifik terhadap dokumen.

1. **Naive Bayes + Bags-of-Words (BoW)**

BoW menghitung frekuensi kemunculan kata tanpa memperhatikan konteks. Meski sederhana, metode ini tetap efekti dalam berbagai aplikasi klasifikasi teks.

1. **Naive Bayes + Word2Vec**

Word2Vec menghasilkan representasi kata dalam bentuk vektor yang menangkap hubungan semantik antar kata. Dalam penelitian ini, vektor tiap kata dirata-ratakan untuk membentuk representasi dokumen.

1. **Naive Bayes + TF-IDF + Feature Selection (Chi-Squared)**

Pendekatan ini Menggunakan TF-IDF sebagai dasar representasi teks, kemudian diterapkan metode seleksi fitur meggunakan uji Chi-Squared untuk memilih fitur kata yang paling relevan.

Model Naive Bayes yang dilatih dengan keempat pendekatan tersebut kemudian dievaluasi Menggunakan beberapa metrik untuk mengetahui efektivitas masing-masing metode dalam mengklasifikasikan sentimen ulasan.

## **Evaluasi Kinerja**

Untuk mengevaluasi model Naive Bayes yang dilatih dengan berbagai metode representasi teks, dilakukan pengukuran menggunakan beberapa metrik evaluasi standar dalam klasifikasi teks, yaitu:

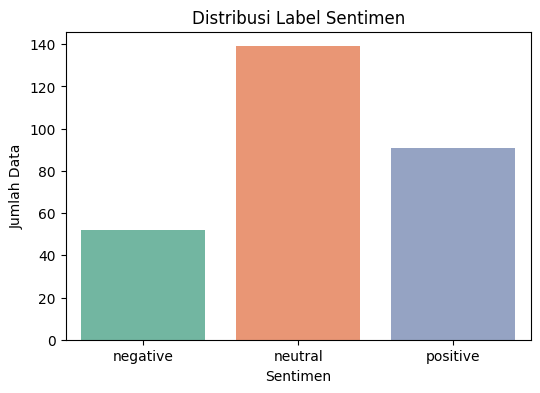
* **Akurasi:** Mengukur proporsi prediksi yang benar terhadap seluruh data uji.
* **Presisi:** Mengukur seberapa tepat model dalam memprediksi label tertentu, khususnya untuk kelas minoritas seperti “positif” dan “negatif”.
* **Recall:** Mengukur seberapa baik model dalam menemukan semua instance dari kelas tertentu.
* **F1-Score:** Rata-rata harmonis dari presisi dan recall.
* **Confusion Matrix:** Matriks evaluasi yang menampilkan jumlah prediksi benar dan salah untuk masing-masing kelas sentimen (positif, negatif, netral).

# BAB IV HASIL DAN PEMBAHASAN

## **Visualisasi dan EDA**

Eksplorasi data dilakukan untuk memahami karakteristik awal dataset ulasan Gojek yang diperoleh dari platform X (Twitter).

### Distribusi Label Sentimen



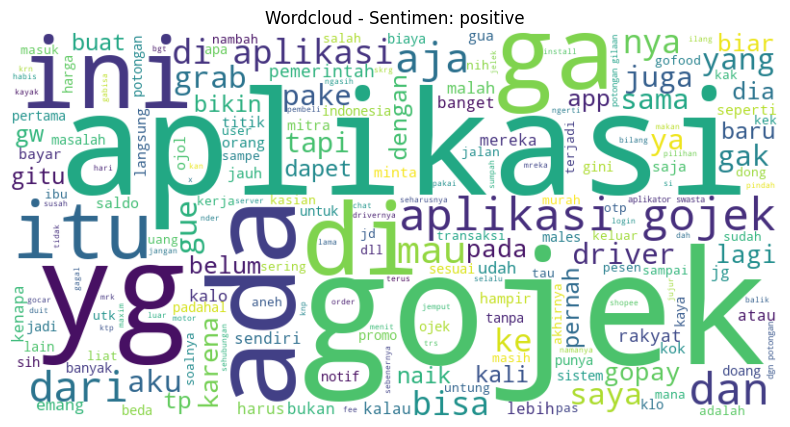
Dari total 282data awal, setelah dilakukan preprocessing, hanya **282** data yang layak digunakan. Distribusi sentimen setelah pelabelan menggunakan model Transformer (seperti BERT) adalah sebagai berikut:

* + 1. **Positif : 91** (≈32.3%)
    2. **Netral :** 139(≈49.3%)
    3. **Negatif :** 52 (≈18.4%)

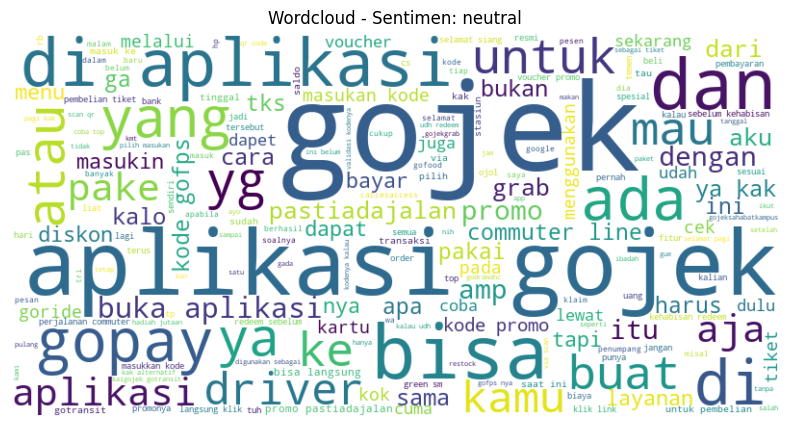
### Word Cloud

WordCloud dibuat untuk masing-masing label sentimen:

1. Positif : kata-kata yang dominan seperti “aplikasi”, “gojek”, “yg”, dan “bisa” mendominasi



1. Netral : kata-kata seperti “untuk”, “ada”, “pake”, dan “itu” kata yang muncul paling sering



1. Negatif : kata-kata seperti “gak”, “pastiadadijalan”, “terus”, dan “buat” kata yang sering muncul.



### Statistik Deskriptif

Analisis statistic dasar yang dilakukan pada Panjang ulasan dan frekuensi kata

1. Rata-rata Panjang ulasan : 27 kata
2. Maksimal Panjang ulasan : 51 kata
3. Minimal Panjang ulasan : 3 kata

## **Hasil *Preprocessing* dan *Feature Engineering***

Preprocessing dilakukan dengan menggunakan tahap case folding, cleansing, tokenizing, dan stopword removal

Contoh hasil preprocessing

1. Sebelum : “Pesan Taksi Listrik Green SM Kini Bisa via Aplikasi Gojek”
2. Sesudah : “pesan taksi listrik green sm kini bisa via aplikasi gojek”

Setelah dilakukan preprocessing dan pelabelan data, selanjutnya fitur teks direpresentasikan ke dalam bentuk numerik menggunakan tiga pendekatan utama :

1. TF-IDF : digunakan sebagai representasi utama karena mampu menekankan kata-kata yang memililki makna penting dalam suatu dokumen.
2. Bag of Words (Bow) : digunakan sebagai pembanding dengan hanya mempertimbangkan frekuensi kemunculan kata tanpa mempertimbangkan konteks global
3. Word2Vec Pretrained Embedding : Untuk memahami makna semantik dari kata-kata, digunakan representasi berbasis Word2Vec pralatih (GloVe dengan 100 dimensi). Setiap kalimat direpresentasikan sebagai rata-rata dari vektor semua kata yang dikenali oleh model tersebut.

Untuk mengurangi risiko overfitting, dilakukan seleksi fitur menggunakan metode Chi-Square(x2). Dari total 5000 fitur TF-IDF, hanya 1000 fitur paling relevan yang dipilih berdasarkan korelasinya dengan label kelas. Seleksi ini membantu mempercepat proses pelatihan dan meningkatkan generalisasi model.

## **Hasil *Preprocessing* dan Pemodelan**

Model *Naïve Bayes* dievaluasi dengan menggunakan pembagian data train-test sebesar 80:20. Eksperimen dilakukan pada tiga jenis representasi fitur teks, yaitu TF-IDF, Bag of Words (BoW), dan embedding Word2Vec pralatih. Selain itu, dilakukan pengujian tambahan dengan menerapkan seleksi fitur menggunakan metode Chi-Square pada representasi TF-IDF untuk mereduksi dimensi dan meningkatkan efisiensi model.

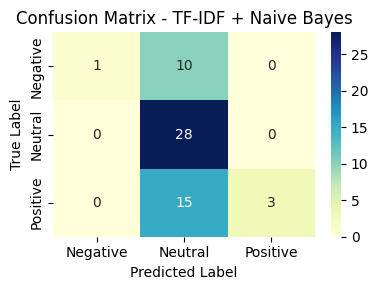
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Accuracy | Precision | Recall | F1-Score |
| Naive Bayes + TF-IDF | 0.56 | 0.77 | 0.56 | 0.46 |
| Naive Bayes + BoW | 0.70 | 0.72 | 0.70 | 0.68 |
| Naive Bayes + Word2Vec | 0.51 | 0.57 | 0.51 | 0.52 |
| Naive Bayes + TF-IDF+Chi2 | 0.53 | 0.57 | 0.53 | 0.39 |

Berdasarkan Tabel, model Naive Bayes dengan representasi Bag of Words (BoW) menunjukkan performa terbaik, dengan akurasi 0.70 dan F1-Score 0.68. Ini mengindikasikan bahwa pendekatan berbasis frekuensi kata sederhana lebih stabil dibanding metode lainnya untuk data ini.

Sementara itu, penggunaan TF-IDF, meskipun umum digunakan, menghasilkan performa lebih rendah (akurasi 0.56, F1-Score 0.46), kemungkinan karena distribusi kata yang tidak merata dan dominasi kelas tertentu.

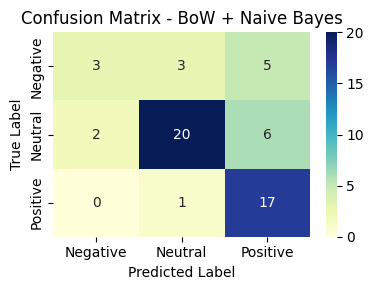
Model dengan Word2Vec pretrained menunjukkan akurasi 0.51 dan F1-Score 0.52, diduga karena metode rata-rata vektor kehilangan konteks antar kata meskipun tetap menangkap makna semantik. Adapun penerapan Chi-Square untuk seleksi fitur pada TF-IDF justru menurunkan performa (akurasi 0.53, F1-Score 0.39), menandakan bahwa reduksi dimensi dalam kasus ini menyebabkan hilangnya informasi penting untuk klasifikasi sentimen.

### Confusion Matrix



Model TF-IDF + Naive Bayes menunjukkan performa yang sangat baik dalam mengenali kelas netral. Dari total 28 data netral, seluruhnya berhasil diklasifikasikan dengan benar tanpa ada kesalahan prediksi ke kelas lain. Untuk kelas negatif, terdapat 11 data, namun hanya 1 data yang diklasifikasikan dengan benar sebagai negatif, sedangkan 10 data lainnya justru diprediksi sebagai netral. Sementara itu, untuk kelas positif (18 data), model ini salah mengklasifikasikan 15 data sebagai netral dan hanya 3 data yang diprediksi dengan benar sebagai positif.

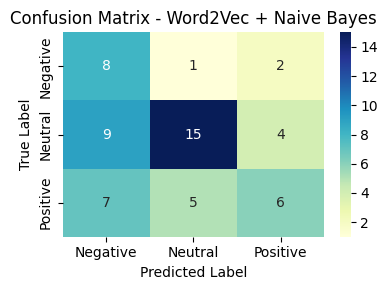
Hal ini mengindikasikan bahwa model TF-IDF + Naive Bayes sangat andal dalam mengenali kelas netral dengan akurasi sempurna, tetapi memiliki keterbatasan signifikan dalam membedakan kelas positif dan negatif. Model cenderung bias terhadap kelas netral, di mana sebagian besar data dari kelas minoritas (positif dan negatif) terserap ke kelas mayoritas (netral).



Model BoW + Naive Bayes menunjukkan performa yang lebih seimbang dibandingkan metode sebelumnya. Untuk kelas netral, dari 28 data berhasil mengklasifikasikan 20 data dengan benar, namun 2 data salah diprediksi sebagai negatif dan 6 data lainnya salah diprediksi sebagai positif.

Pada kelas positif, model menunjukkan performa yang baik dengan mengidentifikasi 17 dari 18 data dengan benar, hanya 1 data yang salah diklasifikasi sebagai netral. Sementara untuk kelas negatif, dari 11 data total, model berhasil mengklasifikasikan 3 data dengan benar, namun 3 data lainnya salah diprediksi sebagai netral dan 5 data salah diprediksi sebagai positif.

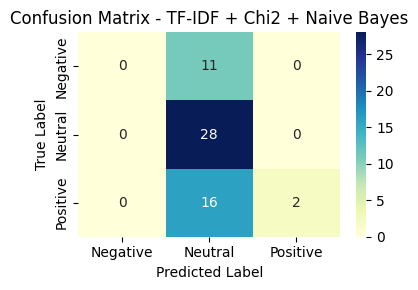
Secara keseluruhan, model BoW + Naive Bayes menunjukkan kemampuan yang relatif baik dalam mengenali kelas positif dan cukup baik untuk kelas netral, tetapi masih mengalami kesulitan dalam mengidentifikasi kelas negatif yang cenderung terdistribusi ke kelas lainnya.



Model Word2Vec + Naive Bayes menunjukkan performa yang kurang optimal dengan distribusi prediksi yang cukup tersebar. Untuk kelas netral, dari 28 data hanya berhasil mengklasifikasikan 15 data dengan benar, sementara 9 data salah diprediksi sebagai negatif dan 4 data salah diprediksi sebagai positif.

Pada kelas negatif, dari 11 data total, model mengidentifikasi 8 data dengan benar, namun 1 data salah diklasifikasi sebagai netral dan 2 data salah diprediksi sebagai positif. Untuk kelas positif, performa model paling lemah dengan hanya 6 dari 18 data yang terklasifikasi benar, sedangkan 7 data salah diprediksi sebagai negatif dan 5 data salah diprediksi sebagai netral.

Secara keseluruhan, model Word2Vec + Naive Bayes menunjukkan kesulitan dalam semua kategori klasifikasi, dengan akurasi yang rendah dan kecenderungan untuk mendistribusikan prediksi secara tidak konsisten ke berbagai kelas, terutama lemah dalam mengenali sentimen positif.



Model TF-IDF + Chi2 + Naive Bayes menunjukkan performa yang sangat baik untuk kelas netral dengan akurasi sempurna, berhasil mengklasifikasikan semua 28 data netral dengan benar tanpa kesalahan prediksi ke kelas lain.

Namun, model mengalami kesulitan signifikan pada kelas lainnya. Untuk kelas negatif, dari 11 data total, seluruhnya salah diklasifikasi sebagai netral tanpa ada satu pun yang teridentifikasi dengan benar. Begitu pula dengan kelas positif, dari 18 data hanya 2 yang berhasil diklasifikasi dengan benar, sementara 16 data lainnya salah diprediksi sebagai netral.

Kondisi ini menunjukkan model memiliki bias yang sangat kuat terhadap kelas netral, hampir tidak mampu mengenali pola sentimen positif dan negatif. Penambahan feature selection Chi2 tampaknya membuat model menjadi terlalu konservatif dalam memprediksi kelas minoritas, sehingga cenderung mengklasifikasikan sebagian besar data sebagai netral.

## **Interpretasi Hasil dan Analisis Komparatif**

Keempat metode kombinasi feature extraction dengan algoritma Naive Bayes yang diuji menunjukkan performa yang bervariasi dalam klasifikasi sentimen. BoW + Naive Bayes memberikan hasil terbaik dengan akurasi keseluruhan yang paling seimbang antar kelas, mampu mengidentifikasi 17 dari 18 data positif dengan benar (94.4% akurasi) dan 20 dari 28 data netral (71.4% akurasi). Meskipun TF-IDF + Naive Bayes dan TF-IDF + Chi2 + Naive Bayes menunjukkan akurasi sempurna untuk kelas netral (100%), kedua model ini memiliki bias yang sangat kuat terhadap kelas mayoritas sehingga hampir tidak mampu mengenali sentimen positif dan negatif.

Word2Vec + Naive Bayes menunjukkan distribusi prediksi yang paling tersebar namun dengan konsistensi yang rendah, terutama lemah dalam mengklasifikasikan sentimen positif (33.3% akurasi). Meskipun model berbasis TF-IDF memiliki precision yang tinggi untuk kelas netral, keseimbangan klasifikasi yang buruk membuat performa keseluruhan menjadi tidak optimal. Oleh karena itu, BoW + Naive Bayes tetap dipilih sebagai model terbaik dalam eksperimen ini karena memberikan trade-off yang paling baik antara akurasi per kelas dan keseimbangan prediksi antar kategori sentimen.

## **Analisis Keunggulan dan Keterbatasan**

Penelitian ini menggabungkan berbagai metode representasi teks (TF-IDF, BoW, Word2Vec, dan TF-IDF + Chi-Square) dengan algoritma Naive Bayes untuk klasifikasi sentimen ulasan berbahasa Indonesia. Pendekatan ini memungkinkan evaluasi menyeluruh terhadap pengaruh fitur terhadap performa model.

Model Naive Bayes + BoW menunjukkan performa terbaik dengan akurasi 70% dan F1-Score 0.68, menandakan efektivitas pendekatan sederhana dalam klasifikasi teks. Sementara TF-IDF dan Word2Vec kurang stabil dan menunjukkan bias terhadap kelas mayoritas atau netral. Chi-Square gagal mempertahankan kualitas prediksi secara konsisten.

Keterbatasan utama terletak pada ketidakseimbangan kelas, terutama kelas negatif, yang menyebabkan rendahnya recall (27.3%) dan menurunnya generalisasi model. Selain itu, belum diterapkannya teknik penyeimbangan data (seperti SMOTE) dan model ensemble menjadi kekurangan yang dapat dieksplorasi pada penelitian selanjutnya.

Secara keseluruhan, pendekatan ini efektif untuk deteksi sentimen secara umum, namun perlu pengembangan lebih lanjut untuk meningkatkan akurasi pada kelas minoritas dan mengurangi bias prediksi.Secara keseluruhan, pendekatan ini efektif untuk deteksi sentimen secara umum, namun perlu pengembangan lebih lanjut untuk meningkatkan akurasi pada kelas minoritas dan mengurangi bias prediksi.

# BAB V KESIMPULAN DAN SARAN

## **Kesimpulan**

Penelitian ini telah berhasil melakukan analisis sentimen terhadap ulasan pengguna aplikasi layanan transportasi Gojek yang diambil dari media sosial X (Twitter), menggunakan algoritma Naive Bayes dengan berbagai pendekatan representasi teks, yaitu TF-IDF, Bag-of-Words (BoW), Word2Vec, serta TF-IDF dengan seleksi fitur Chi-Square. Berdasarkan hasil evaluasi, diperoleh bahwa model Naive Bayes + BoW menunjukkan performa terbaik dengan akurasi 70% dan F1-Score 0.68, serta memberikan keseimbangan prediksi antar kelas yang paling baik.

Sementara itu, model dengan representasi TF-IDF dan Word2Vec menunjukkan keterbatasan dalam membedakan kelas minoritas seperti sentimen negatif dan positif, dengan kecenderungan bias terhadap kelas mayoritas (netral). Penambahan seleksi fitur Chi-Square pada TF-IDF tidak menghasilkan peningkatan kinerja, bahkan menurunkan F1-Score, yang mengindikasikan bahwa pengurangan fitur secara agresif dapat menyebabkan kehilangan informasi penting.

Secara umum, pendekatan yang digunakan dalam penelitian ini efektif dalam menangkap persepsi masyarakat terhadap layanan transportasi online berbasis aplikasi, meskipun masih terdapat tantangan dalam penanganan ketidakseimbangan kelas.

## **Jawaban atas Rumusan Masalah**

1. Bagaimana tingkat akurasi algoritma Naive Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi e-commerce layanan transportasi?  
   Algoritma Naive Bayes mampu mengklasifikasikan sentimen ulasan dengan tingkat akurasi bervariasi tergantung pada metode representasi fitur yang digunakan. Representasi Bag-of-Words menghasilkan akurasi tertinggi sebesar 70%, sementara TF-IDF dan Word2Vec menghasilkan akurasi yang lebih rendah.
2. Metode representasi teks manakah yang memberikan kinerja terbaik ketika dikombinasikan dengan Naive Bayes?  
   Representasi Bag-of-Words (BoW) menunjukkan performa terbaik secara keseluruhan dalam klasifikasi sentimen, mengungguli TF-IDF, Word2Vec, dan TF-IDF dengan seleksi fitur Chi-Square baik dari segi akurasi maupun F1-Score.
3. Bagaimana distribusi sentimen dari ulasan pengguna dan aspek apa yang mempengaruhi sentimen tersebut?  
   Distribusi sentimen pada dataset terdiri dari 32,3% positif, 49,3% netral, dan 18,4% negatif. Berdasarkan hasil wordcloud dan eksplorasi data, kata-kata yang paling dominan dalam ulasan positif menunjukkan kepuasan terhadap fitur dan kemudahan layanan, sementara ulasan negatif lebih banyak menyoroti kendala teknis, kegagalan transaksi, dan pengalaman pengguna yang kurang memuaskan.

## **Saran**

1. Pengembangan model:

Penelitian selanjutnya disarankan untuk mengeksplorasi algoritma lain seperti Support Vector Machine (SVM), Random Forest, atau pendekatan deep learning (CNN, LSTM) guna meningkatkan performa klasifikasi, khususnya untuk kelas minoritas.

1. Penanganan ketidakseimbangan kelas:

Perlu dilakukan upaya untuk menyeimbangkan distribusi kelas, misalnya dengan teknik oversampling seperti SMOTE atau penyesuaian bobot kelas saat pelatihan model agar prediksi tidak bias terhadap kelas mayoritas.

1. Ekspansi data dan platform:

Penambahan volume data dari periode yang lebih panjang dan integrasi dari platform lain (misalnya Instagram, TikTok, atau ulasan Google Play Store) dapat memperluas cakupan analisis dan menghasilkan insight yang lebih komprehensif.

# DAFTAR PUSTAKA

[1] V. Baskov and E. Isaeva, “Information Support of Efficiency of the Transportation Process,” *MATEC Web Conf.*, vol. 334, p. 01035, 2021, doi: 10.1051/matecconf/202133401035.

[2] S. K. Djawa, “The Impact of Information & Communication Technology on Land Transportation Service Business on Indonesia (Case Study in Central Sulawesi Province),” *J. Soc. Res.*, vol. 2, no. 5, pp. 1660–1665, 2023, doi: 10.55324/josr.v2i5.853.

[3] T. Yang, C. Gao, J. Zang, D. Lo, and M. Lyu, *TOUR: Dynamic Topic and Sentiment Analysis of User Reviews for Assisting App Release*, vol. 1, no. 03-06–21. Association for Computing Machinery, 2021. doi: 10.1145/3442442.3458612.

[4] E. L. Funnell, B. Spadaro, N. Martin-Key, T. Metcalfe, and S. Bahn, “mHealth Solutions for Mental Health Screening and Diagnosis: A Review of App User Perspectives Using Sentiment and Thematic Analysis,” *Front. Psychiatry*, vol. 13, no. April, pp. 1–17, 2022, doi: 10.3389/fpsyt.2022.857304.

[5] D. Singh, “Analysis Of Public Sentiment of Covid-19 Pandemic, Vaccines, And Lockdowns,” 2022, [Online]. Available: https://scholarworks.sjsu.edu/etd\_projects/1107

[6] D. Papakyriakou and I. S. Barbounakis, “Data Mining Methods: A Review,” *Int. J. Comput. Appl.*, vol. 183, no. 48, pp. 5–19, 2022, doi: 10.5120/ijca2022921884.

[7] L. Efrizoni, S. Defit, M. Tajuddin, and A. Anggrawan, “Komparasi Ekstraksi Fitur dalam Klasifikasi Teks Multilabel Menggunakan Algoritma Machine Learning,” *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 3, pp. 653–666, 2022, doi: 10.30812/matrik.v21i3.1851.

[8] P. R. Saragih, Y. Sibaroni, and S. S. Prasetyowati, “Multi-Aspect Sentiment Analysis on Gojek Application Reviews Using CNN-LSTM Method,” vol. 7, no. 1, 2025, doi: 10.47065/bits.v7i1.7511.

[9] A. R. Sembiring and C. K. Dewa, “Sentiment Analysis On Indonesian Tweets about the 2024 Election,” *Sinkron*, vol. 9, no. 1, pp. 413–422, 2025, doi: 10.33395/sinkron.v9i1.14481.

[10] M. Iqrom, M. Afdal, Rice Novita, Medyantiwi Rahmawita, and Tengku Khairil Ahsyar, “Sentiment Analysis of Gojek, Grab, Maxim Applications Using Support Vector Machine Algorithm,” *INOVTEK Polbeng - Seri Inform.*, vol. 10, no. 1, pp. 237–248, 2025, doi: 10.35314/52fycr56.

[11] A. Amalia, D. Gunawan, and K. Nasution, “Sentiment analysis of GO-JEK services quality using Multi-Label Classification,” *J. Phys. Conf. Ser.*, vol. 1830, no. 1, pp. 0–7, 2021, doi: 10.1088/1742-6596/1830/1/012003.

[12] A. Wahyudi, Arief Tirtana, and Lady Diana Langoy, “Analysis of Gojek’s Brand Perception Utilizing Twitter Hashtag: Sentiment Analysis Using Ekman’s Classification,” *Open Access Indones. J. Soc. Sci.*, vol. 6, no. 2, pp. 927–935, 2023, doi: 10.37275/oaijss.v6i2.153.