

**PERBANDINGAN ANALISIS SENTIMEN KOMENTAR  
MEDIA SOSIAL DAN DATA ETLE DALAM KLASIFIKASI  
KONDISI LALU LINTAS DI KOTA PALEMBANG  
MENGUNAKAN ALGORITMA *NAIVE BAYES***

**SKRIPSI**

**Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer**



**OLEH:**

**HAURA AHNAF  
09011182025011**

**JURUSAN SISTEM KOMPUTER  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA**

**2025**

# HALAMAN PENGESAHAN

## SKRIPSI

### PERBANDINGAN ANALISIS SENTIMEN KOMENTAR MEDIA SOSIAL DAN DATA ETLE DALAM KLASIFIKASI KONDISI LALU LINTAS DI KOTA PALEMBANG MENGUNAKAN ALGORITMA NAIVE BAYES

Sebagai salah satu syarat untuk penyelesaian studi di  
Program Studi S1 Sistem Komputer

Oleh:

**HAURA AHNAF**

**09011182025011**

**Pembimbing 1**

: **Ahmad Fali Oklilas, M. T.**

**NIP. 197210151999031001**

**Mengetahui**

**Ketua Jurusan Sistem Komputer**



**Dr. Ir. Sukemi, M.T**

**196612032006041001**

## HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Jum'at

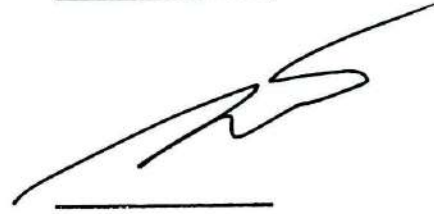
Tanggal : 14 Februari 2025

Tim Penguji:

1. Ketua : Dr. Ir. Sukemi, M.T.



2. Penguji : Dr. Rossi Passarella, M.Eng.



3. Pembimbing : Ahmad Fali Oklilas, M.T.



Mengetahui, 28/2/25

Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M.T.

NIP 196612032006 041001

## HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Haura Ahnaf  
NIM : 09011182025011  
Judul : Perbandingan Analisis Sentimen Komentar Media Sosial dan Data  
ETLE dalam Klasifikasi Kondisi Lalu Lintas di Kota Palembang  
Menggunakan Algoritma Naive Bayes

Hasil Pengecekan Plagiat/Turnitin: 13%

Menyatakan bahwa laporan tugas akhir ini adalah hasil karya saya sendiri dan tidak mengandung unsur penjiplakan atau plagiat. Saya sepenuhnya menyadari bahwa jika terbukti adanya penjiplakan atau plagiat dalam laporan tugas akhir ini, ssaya siap menerima sanksi akademik dan Universitas Sriwijaya.

Demikian pernyataan ini saya buat dengan kesadaran penuh dan tanpa adanya paksaan dari pihak manapun.



Palembang, 24 Maret 2025

Yang Menyatakan,



Haura Ahnaf

NIM 09011182025011

## KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji dan syukur atas kehadiran Allah Subhanahu Wa Ta'ala yang telah melimpahkan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan judul **“Perbandingan Analisis Sentimen Komentar Media Sosial dan Data ETLT dalam Klasifikasi Kondisi Lalu Lintas di Kota Palembang Menggunakan Algoritma Naive Bayes”** yang diajukan sebagai syarat untuk memperoleh gelar sarjana Sistem Komputer Universitas Sriwijaya.

Selesainya penulisan skripsi ini tidak terlepas dari peran semua pihak. Oleh karena itu, pada kesempatan kali ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah Subhanahu Wata'ala yang telah memberikan berkah serta nikmat dan kesehatan kepada penulis dalam melaksanakan penulisan skripsi ini.
2. Orang tua, saudara, dan keluarga besar yang telah memberikan doa dan dukungannya kepada penulis.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya
4. Bapak Dr. Ir. Sukemi, M.T. selaku Ketua Jurusan Sistem Komputer.
5. Bapak Dr. Ahmad Zarkasi, M.T. selaku Dosen Pembimbing Akademik.
6. Bapak Ahmad Fali Oklilas, M.T. selaku Dosen Pembimbing Tugas Akhir atas saran, masukan, dan ilmu yang bermanfaat guna kelancaran penulisan skripsi.
7. Kak Angga selaku Admin Jurusan Sistem Komputer yang telah membantu melancarkan proses administrasi.
8. Dosen-dosen jurusan Sistem Komputer Universitas Sriwijaya yang telah memberikan ilmu yang bermanfaat sehingga menjadi bekal dalam penulisan skripsi.
9. Teman-teman seperjuangan yang telah memberikan dukungannya selama penyelesaian skripsi.
10. Semua pihak yang telah membantu.

Penulis menyadari dalam penyusunan skripsi ini masih terdapat banyak kekurangan. Untuk itu, penulis mengharapkan kritik dan saran yang bersifat membangun dari semua pihak agar skripsi ini dapat segera diperbaiki dan menjadi lebih baik lagi. Penulis berharap agar skripsi ini dapat menjadi ilmu yang bermanfaat bagi semua pembacanya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Palembang, 24 Maret 2025

Penulis,



Haura Ahnaf

NIM. 09011182025011

# **PERBANDINGAN ANALISIS SENTIMEN KOMENTAR MEDIA SOSIAL DAN DATA ETLE DALAM KLASIFIKASI KONDISI LALU LINTAS DI KOTA PALEMBANG MENGUNAKAN ALGORITMA NAIVE BAYES**

**HAURA AHNAF (09011182025011)**

Jurusan Sistem Komputer, Fakultas Ilmu Komputer,

Universitas Sriwijaya

Email: [hauraahnaf01@gmail.com](mailto:hauraahnaf01@gmail.com)

## **ABSTRAK**

Manajemen lalu lintas di kota besar seperti Palembang menghadapi tantangan kepadatan, terutama di jalan protokol. Penerapan ETLE menyediakan data real-time mengenai kondisi lalu lintas, termasuk jumlah kendaraan, untuk membantu pengelolaan arus kendaraan. Sementara itu, media sosial menjadi sumber alternatif informasi melalui laporan pengguna mengenai kemacetan dan kecelakaan. Penelitian ini bertujuan untuk membandingkan data dari media sosial terhadap data lalu lintas dari ETLE menggunakan algoritma naive bayes. Dari hasil penelitian, didapatkan akurasi training 97,99%, validation 97,21%, dan testing sebesar 95,77% pada data media sosial. Sedangkan pada data ETLE didapatkan akurasi training 84,4%, validation 81,6%, dan testing sebesar 95,17%. Dari hasil klasifikasi kedua data, didapatkan persentase kemiripan hasil klasifikasi sebesar 77,93%. Selain itu, hasil penelitian menunjukkan bahwa 50,8% dari data menunjukkan sentimen negatif terhadap kondisi lalu lintas melalui media sosial. Penelitian berikutnya diharapkan agar dataset lebih seimbang guna menghindari ketimpangan data serta mempertimbangkan penggunaan algoritma yang lebih kompleks untuk menangani permasalahan lalu lintas secara lebih mendalam.

**Kata Kunci** : *Naive Bayes*, *Sentiment Analysis*, Lalu Lintas, ETLE, Kota Palembang

**COMPARISON OF SENTIMENT ANALYSIS OF SOCIAL  
MEDIA COMMENTS AND ETLE DATA IN CLASSIFYING  
TRAFFIC CONDITIONS IN PALEMBANG CITY USING THE  
NAIVE BAYES ALGORITHM**

**HAURA AHNAF (09011182025011)**

*Department of Computer Systems, Faculty of Computer Science,  
Sriwijaya University*

Email: [hauraahnaf01@gmail.com](mailto:hauraahnaf01@gmail.com)

**ABSTRACT**

*Traffic management in major cities like Palembang faces congestion challenges, especially on main roads. The implementation of ETLE provides real-time data on traffic conditions, including vehicle counts, to assist in traffic flow management. Meanwhile, social media serves as an alternative source of information through user reports on congestion and accidents. This study aims to compare data from social media with traffic data from ETLE using the Naïve Bayes algorithm. The results show that the training accuracy for social media data reached 97,99%, with validation accuracy of 97,21%, and testing accuracy of 95,77%. Meanwhile, for ETLE data, the training accuracy was 84,4%, with validation accuracy of 81.6%, and testing accuracy of 95,17%. The classification results indicate a similarity percentage of 77.93% between the two datasets. Additionally, the study found that 50,8% of the data from social media expressed negative sentiment toward traffic conditions. Future research is expected to use a more balanced dataset to avoid data imbalance and consider employing more advanced algorithms to address traffic issues more comprehensively.*

**Keywords:** *Naïve Bayes, Sentiment Analysis, Traffic, ETLE, Palembang City*



## DAFTAR ISI

<b>HALAMAN PENGESAHAN</b> .....	<b>ii</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>iii</b>
<b>HALAMAN PERNYATAAN</b> .....	<b>iv</b>
<b>KATA PENGANTAR</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>vii</b>
<b>ABSTRACT</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>ix</b>
<b>DAFTAR GAMBAR</b> .....	<b>xii</b>
<b>DAFTAR TABEL</b> .....	<b>xiii</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xiv</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah .....	2
1.3. Batasan Masalah .....	3
1.4. Tujuan.....	3
1.5. Manfaat.....	3
1.6. Metodologi Penelitian .....	4
1.7. Sistematika Penulisan .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>7</b>
2.1. Penelitian Terkait.....	7
2.2. Data ETL.....	14
2.3. Data Media Sosial.....	14
2.4. Kepadatan Lalu Lintas.....	15
2.5. Jalan Protokol di Kota Palembang.....	16
2.6. <i>Naive Bayes</i> .....	17
2.7. <i>Web Scraping</i> .....	18
2.8. Analisis Sentimen .....	19
2.9. Prapemrosesan Data .....	20
2.9.1. <i>Cleaning Data</i> .....	20
2.9.2. <i>Case Folding</i> .....	20
2.9.3. <i>Stemming</i> .....	20
2.9.4. <i>Tokenization</i> .....	20
2.9.5. <i>Stopword Removal</i> .....	21
2.9.6. <i>Normalization</i> .....	21
2.10. Pembobotan ( <i>Term Weighting</i> ).....	21
2.11. <i>Confusion Matrix</i> .....	22
2.11.1. <i>Accuracy</i> .....	23
2.11.2. <i>Recall</i> .....	23
2.11.3. <i>Precision</i> .....	24
2.11.4. <i>F1-Score</i> .....	24
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>25</b>
3.1. Alur Penelitian.....	25
3.2. Menentukan Topik Penelitian.....	26

3.3.	Identifikasi Masalah .....	27
3.4.	Studi Pustaka dan Literatur.....	28
3.5.	Kebutuhan Perangkat.....	28
3.6.	Pengumpulan Data.....	29
3.6.1.	Data Media Sosial.....	29
3.6.2.	Data Angka Kendaraan.....	30
3.6.3.	Data Tabel Referensi .....	32
3.7.	Prapemrosesan Data .....	33
3.7.1.	<i>Cleaning Data</i> .....	33
3.7.2.	<i>Case Folding</i> .....	34
3.7.3.	<i>Stemming</i> .....	34
3.7.4.	<i>Tokenization</i> .....	35
3.7.5.	<i>Stopwords Removal</i> .....	35
3.7.6.	<i>Normalization</i> .....	36
3.8.	Membagi Data Menjadi Data Latih dan Data Uji.....	36
3.9.	Pembobotan Data.....	37
3.9.1.	Menghitung Nilai TF .....	37
3.9.2.	Menghitung Nilai IDF .....	38
3.9.3.	Menghitung Nilai TF-IDF .....	38
3.10.	Klasifikasi Data Menggunakan <i>Naïve Bayes</i> .....	38
3.10.1.	Klasifikasi Data Teks.....	38
3.10.2.	Klasifikasi Data Angka Kendaraan .....	39
3.10.3.	Output <i>Naïve Bayes</i> .....	40
3.11.	Perbandingan Hasil Klasifikasi Data Teks dengan Data ETLE .....	40
3.12.	Evaluasi Model.....	41
3.13.	Analisis .....	41
3.14.	Kesimpulan dan Saran .....	41
<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>42</b>
4.1.	Pengumpulan Data Teks.....	42
4.2.	<i>Preprocessing Data</i> .....	45
4.3.	<i>Data Cleaning</i> .....	46
4.4.	<i>Case Folding</i> .....	47
4.5.	<i>Stemming</i> .....	47
4.6.	<i>Tokenization</i> .....	48
4.7.	<i>Stopword Removal</i> .....	49
4.8.	<i>Normalization</i> .....	50
4.9.	<i>Filtering Data</i> .....	51
4.10.	Penilaian Sentimen Terhadap Data Media Sosial .....	51
4.11.	<i>Splitting Data</i> .....	52
4.12.	Pembobotan TF-IDF.....	53
4.13.	Evaluasi Model Klasifikasi Teks.....	57
4.14.	Klasifikasi Menggunakan Model Teks <i>Naïve Bayes</i> .....	59
4.15.	Visualisasi Data Teks .....	61
4.16.	Evaluasi Model Klasifikasi ETLE.....	62
4.17.	Klasifikasi Menggunakan Model ETLE <i>Naïve Bayes</i> .....	65
4.18.	Perbandingan Data Klasifikasi Media Sosial dan ETLE.....	66

4.19. Analisis .....	69
<b>BAB V KESIMPULAN .....</b>	<b>72</b>
5.1. Kesimpulan.....	72
5.2. Saran .....	72
<b>DAFTAR PUSTAKA .....</b>	<b>74</b>
<b>LAMPIRAN.....</b>	<b>80</b>

## DAFTAR GAMBAR

<b>Gambar 2. 1.</b> Analisis Sentimen .....	19
<b>Gambar 3. 1.</b> Alur Penelitian .....	25
<b>Gambar 3. 2.</b> Data Teks Instagram dan Tiktok.....	30
<b>Gambar 3. 3.</b> Data Sebelum dan Sesudah Dibersihkan .....	34
<b>Gambar 3. 4.</b> Data Sebelum dan Sesudah <i>Case Folding</i> .....	34
<b>Gambar 3. 5.</b> Data Sebelum dan Sesudah <i>Stemming</i> .....	35
<b>Gambar 3. 6.</b> Data Sebelum dan Sesudah Tokenisasi .....	35
<b>Gambar 3. 7.</b> Data Sebelum dan Sesudah <i>Stopwords Removal</i> .....	36
<b>Gambar 3. 8.</b> Data Sebelum dan Sesudah Normalisasi .....	36
<b>Gambar 3. 9.</b> Hasil Pembobotan TF .....	37
<b>Gambar 3. 10.</b> Hasil Pembobotan IDF .....	38
<b>Gambar 3. 11.</b> Hasil Pembobotan TF-IDF .....	38
<b>Gambar 4. 1.</b> URL Instagram dan Tiktok.....	42
<b>Gambar 4. 2.</b> Pengambilan Data Instagram Pada Apify.....	43
<b>Gambar 4. 3.</b> Pengambilan Data Tiktok Pada Apify .....	43
<b>Gambar 4. 4.</b> Pengambilan Data IG Dengan Ektensi .....	44
<b>Gambar 4. 5.</b> <i>Dataset</i> Instagram.....	44
<b>Gambar 4. 6.</b> <i>Dataset</i> Tiktok .....	45
<b>Gambar 4. 7.</b> <i>Dataset</i> Instagram dan Tiktok .....	45
<b>Gambar 4. 8.</b> <i>Cleaning</i> Data Teks .....	46
<b>Gambar 4. 9.</b> <i>Case folding</i> Data Teks.....	47
<b>Gambar 4. 10.</b> <i>Stemming</i> Data Teks .....	48
<b>Gambar 4. 11.</b> <i>Tokenization</i> Data Teks .....	48
<b>Gambar 4. 12.</b> <i>Stopword Removal</i> Data Teks.....	49
<b>Gambar 4. 13.</b> <i>Normalization</i> Data Teks.....	50
<b>Gambar 4. 14.</b> <i>Filtering</i> Data Teks.....	51
<b>Gambar 4. 15.</b> Skor Polaritas Media Sosial.....	52
<b>Gambar 4. 16.</b> Data Latih Teks.....	53
<b>Gambar 4. 17.</b> Data Uji Teks.....	53
<b>Gambar 4. 18.</b> Pembobotan Data Teks .....	55
<b>Gambar 4. 19.</b> Nilai TF-IDF Tertinggi .....	56
<b>Gambar 4. 20.</b> Evaluasi Model <i>Naïve Bayes</i> Terhadap Data Teks.....	57
<b>Gambar 4. 21.</b> Visualisasi Data Teks .....	62
<b>Gambar 4. 22.</b> Evaluasi Model <i>Naïve Bayes</i> Terhadap Data ETLE.....	63

## DAFTAR TABEL

<b>Tabel 2. 1.</b> Daftar Penelitian Terkait .....	7
<b>Tabel 2. 2.</b> Daftar Jalan Protokol di Kota Palembang .....	16
<b>Tabel 2. 3.</b> <i>Confusion Matrix</i> 2x2.....	23
<b>Tabel 3. 1.</b> Spesifikasi Perangkat <i>Hardware</i> .....	28
<b>Tabel 3. 2.</b> Spesifikasi Perangkat <i>Software</i> .....	29
<b>Tabel 3. 3.</b> Data Jumlah Kendaraan ETLE.....	30
<b>Tabel 3. 4.</b> Nilai Input Panjang Jalan .....	31
<b>Tabel 3. 5.</b> Data Tabel Referensi .....	32
<b>Tabel 3. 6.</b> Tabel Konversi Data Referensi .....	33
<b>Tabel 3. 7.</b> Input Data Angka Kendaraan .....	39
<b>Tabel 3. 8.</b> <i>Output Naïve Bayes</i> .....	40
<b>Tabel 4. 1.</b> Penilaian Sentimen Media Sosial .....	52
<b>Tabel 4. 2.</b> Contoh Data Teks .....	55
<b>Tabel 4. 3.</b> Perhitungan <i>Confusion Matrix</i> Model Klasifikasi Data Teks .....	58
<b>Tabel 4. 4.</b> Perhitungan <i>Precision</i> Model Klasifikasi Data Teks .....	58
<b>Tabel 4. 5.</b> Perhitungan <i>Recall</i> Model Klasifikasi Data Teks .....	59
<b>Tabel 4. 6.</b> Perhitungan <i>F1-Score</i> Model Klasifikasi Data Teks.....	59
<b>Tabel 4. 7.</b> Hasil Klasifikasi Model Teks <i>Naïve Bayes</i> .....	60
<b>Tabel 4. 8.</b> Perhitungan <i>Confusion Matrix</i> Model Klasifikasi Data ETLE .....	63
<b>Tabel 4. 9.</b> Perhitungan <i>Precision</i> Model Klasifikasi Data ETLE .....	64
<b>Tabel 4. 10.</b> Perhitungan <i>Recall</i> Model Klasifikasi Data ETLE .....	64
<b>Tabel 4. 11.</b> Perhitungan <i>F1-Score</i> Model Klasifikasi Data ETLE.....	65
<b>Tabel 4. 12.</b> Hasil Klasifikasi Model ETLE <i>Naïve Bayes</i> .....	65
<b>Tabel 4. 13.</b> Perbandingan Data Media Sosial dan ETLE.....	67

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b>	Form Perbaikan Penguji.....	L1
<b>Lampiran 2.</b>	Form Perbaikan Pembimbing .....	L2
<b>Lampiran 3.</b>	Hasil <i>Similarity Check</i> .....	L3
<b>Lampiran 4.</b>	URL Postingan Media Sosial.....	L5
<b>Lampiran 5.</b>	Data Mentah.....	L6
<b>Lampiran 6.</b>	<i>Preprocessing Data</i> .....	L7
<b>Lampiran 7.</b>	<i>Filtering Data</i> .....	L8
<b>Lampiran 8.</b>	Penilaian Sentimen .....	L9
<b>Lampiran 9.</b>	Data Latih .....	L10
<b>Lampiran 10.</b>	Data Uji.....	L11
<b>Lampiran 11.</b>	Hasil TF-IDF .....	L12
<b>Lampiran 12.</b>	Hasil Klasifikasi Media Sosial.....	L13
<b>Lampiran 13.</b>	Tabel Referensi .....	L14
<b>Lampiran 14.</b>	Data Jumlah Kendaraan .....	L15
<b>Lampiran 15.</b>	Hasil Klasifikasi ETLE.....	L16
<b>Lampiran 16.</b>	Hasil Perbandingan Klasifikasi Media Sosial dan ETLE .....	L17

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Manajemen lalu lintas adalah aspek penting dalam pengelolaan kota-kota besar. Kepadatan lalu lintas tidak hanya mengganggu kelancaran mobilitas warga, tetapi juga berdampak signifikan terhadap ekonomi, lingkungan, dan kualitas hidup masyarakat [1]. Kota Palembang sebagai salah satu kota besar di Indonesia tidak lepas dari masalah kepadatan lalu lintas, terutama di jalan-jalan protokol yang menjadi jalur utama pergerakan warga.

Dalam upaya mengatasi permasalahan tersebut, Pemerintah Kota Palembang telah mengimplementasikan sistem Electronic Traffic Law Enforcement (ETLE) di sejumlah titik strategis. ETLE adalah salah satu teknologi yang diterapkan untuk membantu mengelola dan memantau lalu lintas. ETLE menggunakan kamera dan sensor untuk menangkap pelanggaran lalu lintas secara otomatis dan menyediakan data *real-time* tentang kondisi lalu lintas [2]. Selain itu, ETLE dapat mendeteksi jumlah kendaraan yang melewati kamera untuk berbagai jenis kendaraan.

Di sisi lain, media sosial muncul sebagai sumber informasi alternatif mengenai kondisi lalu lintas. Pengguna media sosial sering kali membagikan informasi real-time tentang kemacetan, kecelakaan, dan kondisi jalan lainnya. Data ini, meskipun tidak selalu terstruktur, dapat dianalisis untuk mendapatkan wawasan yang berharga tentang kondisi lalu lintas [3]. Salah satu metode yang dapat digunakan untuk menganalisis data ini adalah analisis sentimen, yang menilai opini publik dan perasaan mereka terhadap suatu topik, dalam hal ini kondisi lalu lintas.

Meskipun data dari media sosial menawarkan keuntungan dalam hal kecepatan dan luas cakupan, data ini juga memiliki tantangan seperti ketidakakuratan, rentan terhadap bias, informasi yang bersifat subjektif, dan masalah privasi. Pengguna media sosial juga cenderung lebih aktif dalam melaporkan kejadian negatif seperti kemacetan daripada kejadian positif seperti kelancaran lalu lintas, yang dapat menyebabkan distorsi dalam analisis. Oleh karena

itu, penelitian ini bertujuan untuk membandingkan data yang diperoleh dari media sosial dengan data yang dihasilkan oleh sistem ETLE. Sehingga tingkat akurasi dan reliabilitas informasi yang berasal dari media sosial dapat dievaluasi secara lebih objektif.

Algoritma *Naive Bayes* adalah salah satu metode yang efektif untuk klasifikasi berbasis data. Algoritma ini dapat digunakan untuk memodelkan probabilitas kepadatan lalu lintas berdasarkan data historis dari ETLE dan hasil analisis sentimen dari media sosial [4]. Selain itu, algoritma *Naive Bayes* telah digunakan dalam analisis media sosial untuk mendeteksi peristiwa lalu lintas secara real-time, seperti yang dijelaskan dalam penelitian [5]. Pendekatan ini memungkinkan integrasi data dari berbagai sumber untuk meningkatkan akurasi klasifikasi kondisi lalu lintas.

Penelitian [6] mengklasifikasikan dua kelompok, negatif dan positif untuk melakukan analisis sentimen terhadap data twitter menggunakan topik kemacetan lalu lintas. Setelah mengevaluasi algoritma *Naive Bayes*, hasil tertinggi mencapai akurasi keseluruhan sebesar 77%, presisi sebesar 86%, *recall* sebesar 82%, dan *f1-score* sebesar 84%. Kemudian penelitian [7] yang membandingkan metode *naive bayes* dengan *support vector machine* dalam analisis sentimen twitter memberikan hasil komparasi yang menunjukkan algoritma *naive bayes* memiliki hasil yang lebih baik daripada SVM dengan akurasi sebesar 73,65% sedangkan SVM mendapatkan hasil akurasi sebesar 70,20%.

Dengan memanfaatkan algoritma *Naive Bayes*, penelitian ini diharapkan dapat mengevaluasi sejauh mana media sosial dapat menjadi sumber data alternatif yang andal dalam memberikan informasi terkait kondisi lalu lintas, khususnya di kota Palembang.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang di atas, maka rumusan masalah yang akan dibahas dalam penelitian ini sebagai berikut:



1. Bagaimana tingkat keakuratan model klasifikasi kepadatan lalu lintas di jalan protokol kota Palembang dari data media sosial menggunakan algoritma *naïve bayes*?
2. Bagaimana tingkat keakuratan model klasifikasi kepadatan lalu lintas di jalan protokol kota Palembang dari data ETLE menggunakan algoritma *naïve bayes*?
3. Bagaimana tingkat kemiripan hasil klasifikasi kepadatan lalu lintas antara data media sosial dengan data ETLE?

### **1.3. Batasan Masalah**

Batasan Masalah dari tugas akhir ini adalah sebagai berikut:

1. Penelitian difokuskan pada jalan-jalan protokol di kota Palembang.
2. Algoritma yang digunakan dalam melakukan klasifikasi adalah *naïve bayes*.
3. Data media sosial yang akan dianalisis berasal dari *platform* instagram dan tiktok, dan pada periode waktu yang sama dengan data ETLE.

### **1.4. Tujuan**

Tujuan dari pembuatan tugas akhir ini adalah sebagai berikut:

1. Mengetahui tingkat keakuratan model klasifikasi kepadatan lalu lintas di jalan protokol kota Palembang dari data media sosial menggunakan algoritma *naïve bayes*.
2. Mengetahui tingkat keakuratan model klasifikasi kepadatan lalu lintas di jalan protokol kota Palembang dari data ETLE menggunakan algoritma *naïve bayes*.
3. Mengetahui tingkat kemiripan hasil klasifikasi kepadatan lalu lintas antara data media sosial dengan data ETLE.

### **1.5. Manfaat**

Manfaat dari pembuatan tugas akhir ini adalah sebagai berikut:

1. Menghasilkan model untuk klasifikasi kepadatan lalu lintas di jalan protokol kota Palembang.

2. Menambah pengetahuan tentang cara penerapan algoritma *naive bayes* yang optimal untuk melakukan klasifikasi kepadatan lalu lintas.
3. Membantu instansi berwenang dalam pengambilan kebijakan yang tepat untuk mengatur lalu lintas di jalan protokol kota Palembang.

## 1.6. Metodologi Penelitian

Tahapan-tahapan yang dilakukan dalam proses pembuatan tugas akhir adalah sebagai berikut:

### 1. Studi Pustaka dan Literatur

Pada metode ini, penulis mencari dan mengumpulkan referensi mengenai penelitian yang akan dibahas melalui sumber dari buku, artikel ilmiah, dan internet mengenai analisis sentimen di media sosial menggunakan algoritma *naive bayes*.

### 2. Pengumpulan data

Penulis melakukan pengumpulan data dari *platform* media sosial berupa instagram dan tiktok dan data aktual *Electronic Traffic Law Enforcement* (ETLE) dari instansi kepolisian kota Palembang.

### 3. *Preprocessing* data

Pada metode ini, penulis melakukan *preprocessing data* untuk mempersiapkan data mentah yang telah dikumpulkan agar siap digunakan dalam analisis. Proses *preprocessing* yang dilakukan berupa *cleaning data*, *case folding*, *stemming*, *tokenizing*, dan *stopwords removal*.

### 4. Perancangan model

Pada metode ini, penulis membangun model klasifikasi kepadatan lalu lintas menggunakan algoritma *naive bayes*. Model ini dirancang untuk melakukan klasifikasi kepadatan lalu lintas berdasarkan data yang diperoleh dari media sosial dan data historis dari ETLE.

### 5. Perbandingan kemiripan data hasil klasifikasi data media sosial dengan hasil klasifikasi data ETLE

Dalam tahap ini, penulis melakukan perbandingan kemiripan data hasil klasifikasi data media sosial dengan data angka kendaraan dari ETLE menggunakan program dengan logika sederhana.

#### 6. Analisa dan Kesimpulan

Penulis menganalisa hasil yang didapatkan dari penelitian berupa evaluasi kinerja model, identifikasi faktor-faktor yang mempengaruhi akurasi model, dan interpretasi hasil perbandingan antara data hasil klasifikasi media sosial dan data aktual dari ETLE. Kemudian penulis menarik kesimpulan mengenai efektivitas dan akurasi model *Naive Bayes* dalam klasifikasi kepadatan lalu lintas.

### 1.7. Sistematika Penulisan

Sistematika Penulisan dalam pembuatan proposal skripsi sebagai berikut:

#### **BAB I PENDAHULUAN**

Pendahuluan berisi uraian tentang latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, metodologi penelitian, dan sistematika penulisan terkait penelitian Perbandingan Analisis Sentimen Komentar Media Sosial dan Data ETLE dalam Klasifikasi Kondisi Lalu Lintas di Kota Palembang Menggunakan Algoritma Naive Bayes.

#### **BAB II TINJAUAN PUSTAKA**

Tinjauan Pustaka berisi literatur terkait yang telah dikaji dalam konteks penelitian Perbandingan Analisis Sentimen Komentar Media Sosial dan Data ETLE dalam Klasifikasi Kondisi Lalu Lintas di Kota Palembang Menggunakan Algoritma Naive Bayes.

#### **BAB III METODOLOGI PENELITIAN**

Metodologi Penelitian berisi penjelasan secara detail mengenai pendekatan dan teknik yang digunakan dalam pelaksanaan penelitian Perbandingan Analisis Sentimen Komentar Media Sosial dan Data ETLE dalam Klasifikasi Kondisi Lalu Lintas di Kota Palembang Menggunakan Algoritma Naive Bayes.

#### **BAB IV HASIL DAN PEMBAHASAN**

Hasil dan Pembahasan berisi penjelasan proses serta hasil dari penelitian yang dilakukan mengenai Perbandingan Analisis Sentimen Komentar Media Sosial dan Data ETLE dalam Klasifikasi Kondisi Lalu Lintas di Kota Palembang Menggunakan Algoritma Naive Bayes.

## **BAB V PENUTUP**

Penutup berisi kesimpulan yang diperoleh dari hasil penelitian, dan saran untuk pengembangan sistem pada penelitian Perbandingan Analisis Sentimen Komentar Media Sosial dan Data ETLE dalam Klasifikasi Kondisi Lalu Lintas di Kota Palembang Menggunakan Algoritma Naive Bayes.

## DAFTAR PUSTAKA

- [1] Lirebo Sokido, Daniel (2024) “Measuring the level of urban traffic congestion for sustainable transportation in Addis Ababa, Ethiopia, the cases of selected intersections”, *Frontiers in Sustainable Cities*, 6, pp. 1. doi: 10.3389/frsc.2024.1366932.
- [2] Hidayat, Firman (2023) *Penyidikan terhadap pelanggaran lalu lintas dalam sistem electronic traffic law enforcement dalam perspektif keadilan hukum*. Undergraduate Thesis, Available at: <https://repository.unmerpas.ac.id/375/> (Accessed: 7 August 2024).
- [3] Kresna Putra, Prabu (2020) *Pemantauan Lalu Lintas dan Kondisi Jalan dengan Pendekatan Ekstraksi Informasi dari Media Sosial Twitter = Monitoring Traffic and Road Conditions with Information Extraction Approach from Twitter Social Media*. Undergraduate Thesis, Available at: <https://lib.ui.ac.id/detail?id=20510361> (Accessed: 7 August 2024).
- [4] Rachmat, Antonius and Yuan Lukito (2017) “Deteksi Komentar Spam Bahasa Indonesia Pada Instagram Menggunakan *Naive Bayes*”, *Ultimatics Jurnal Teknik Informatika*, 9(1), pp. 52. doi: 10.31937/ti.v9i1.564.
- [5] Vanjare, N. *et al.* (2022) “Real-Time Citizen Problem Detection From Twitter Data Using *Naive Bayes* Classifier”, *SSRN Electronic Journal* [Preprint]. doi: 10.2139/ssrn.4097217.
- [6] Romindo, Okky Putra Barus and Jefri Junifer Pangaribuan (2024) “Implementasi Metode *Naive Bayes* Classifier Terhadap Klasifikasi Topik Kemacetan Lalu Lintas Indonesia Melalui Tweet”, *Jurnal Media Informatika Budidarma*, 8(2), pp. 1095. doi:10.30865/mib.v8i2.7470.
- [7] Izzul Fikri, M. *et al.* (2020) “Perbandingan Metode *Naive Bayes* dan Support Vector Machine pada Analisis Sentimen Twitter”, *SMATIKA JURNAL: STKI Informatika Jurnal*, 10(2), pp. 75. doi:10.32664/smatika.v10i02.455.
- [8] Romindo, R., Barus, O. P. and Pangaribuan, J. J. (2024) ‘Implementasi Metode *Naive Bayes* Classifier Terhadap Klasifikasi Topik Kemacetan Lalu Lintas Indonesia Melalui Tweet’, *Jurnal Media Informatika Budidarma*, 8(2), pp. 1087-1096. doi: 10.30865/mib.v8i2.7470.

- [9] Saepudin, S., Widiastuti, S. and Irawan, C. (2023) ‘Sentiment Analysis of *Social Media Platform* Reviews Using the *Naïve Bayes* Classifier Algorithm’, *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, 12(2), pp. 236–243. doi: 10.32736/sisfokom.v12i2.1650.
- [10] Hendriansyah and Pramusinto, W. (2023) ‘Analisis Sentimen Masyarakat Terhadap Tilang Elektronik Pada Twitter Dengan Algoritma K-Nearest Analysis of Public Sentiment on Electronic Tilang on Twitter *Social Media* Using Countvectorizer Feature Extraction and K-Nearest Neighbor Algorithm’, *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 2(2), pp. 1094–1101. Available at: <https://senafti.budiluhur.ac.id/index.php/senafti/article/view/1030> (Access ed: 15 August 2024)
- [11] Siswanto, S. *et al.* (2022) ‘The Sentiment Analysis Using *Naïve Bayes* with Lexicon-Based Feature on TikTok Application’, *Jurnal Varian*, 6(1), pp. 89–96. doi: 10.30812/varian.v6i1.2205.
- [12] Seliverstov, Y. *et al.* (2020) ‘Traffic safety evaluation in Northwestern Federal District using sentiment analysis of internet users’ reviews’, *Transportation Research Procedia*, 50, pp. 626–635. doi: 10.1016/j.trpro.2020.10.074.
- [13] Odim, M. O. *et al.* (2020) ‘Exploring the performance characteristics of the *naïve bayes* classifier in the sentiment analysis of an Airline’s *social media* data’, *Advances in Science, Technology and Engineering Systems*, 5(4), pp. 266–272. doi: 10.25046/aj050433.
- [14] Robi Padri, A., Asro, A. and Indra, I. (2023) ‘Classification of Traffic Congestion in Indonesia Using the *Naive Bayes* Classification Method’, *Journal of World Science*, 2(6), pp. 877–888. doi: 10.58344/jws.v2i6.285.
- [15] Silitonga, P. D. P. *et al.* (2023) ‘Comparison of Tiktok User Sentiment Analysis Accuracy with *Naïve Bayes* and Support Vector Machine’, *International Journal of Advanced Trends in Computer Science and Engineering*, 12(1), pp. 11–15. doi: 10.30534/ijatcse/2023/031212023.
- [16] Aulia, Z. N., Jati, G. K. and Santoso, I. (2023) ‘Analisis Sentimen Tanggapanpublic Mengenai E-Tilang Melalui Media Sosial Youtube Menggunakan Algoritma *Naive Bayes*’, *Jurnal IKRA-ITH Informatika*, 7(2),

- pp. 150–156. Available at: <https://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/view/2824/2077> (Accessed: 15 August 2024).
- [17] Beenish, H. *et al.* (2023) ‘Contemporary Study of *Machine Learning* Algorithms for Traffic Density Estimation in Intelligent Transportation Systems’, *IJIST*, 5(4). pp. 592-608. Available at: [https://www.researchgate.net/profile/Imran-Imran-19/publication/376173087\\_Contemporary\\_Study\\_of\\_Machine\\_Learning\\_Algorithms\\_for\\_Traffic\\_Density\\_Estimation\\_in\\_Intelligent\\_Transportation\\_Systems/links/656c10003fa26f66f44a1dc1/Contemporary-Study-of-Machine-L](https://www.researchgate.net/profile/Imran-Imran-19/publication/376173087_Contemporary_Study_of_Machine_Learning_Algorithms_for_Traffic_Density_Estimation_in_Intelligent_Transportation_Systems/links/656c10003fa26f66f44a1dc1/Contemporary-Study-of-Machine-L) (Accessed: 15 August 2024).
- [18] Chen, H. *et al.* (2021) ‘Improved *naive Bayes* classification algorithm for traffic risk management’, *Eurasip Journal on Advances in Signal Processing*, 30. pp. 1-12. doi: 10.1186/s13634-021-00742-6.
- [19] Razali, N. A. M. *et al.* (2021) ‘Gap, techniques and evaluation: traffic flow prediction using *machine learning* and deep learning’, *Journal of Big Data*, 8(1). pp. 1-25. doi: 10.1186/s40537-021-00542-7.
- [20] Ororiseng Leroke, G. and Lall, M. (2021) ‘A Near Real-Time Traffic Congestion Monitoring System Using Sentiment Analysis On Twitter Data’, *International Journal of Scientific & Technology Research*, 10(10), pp. 78–87. Available at: [www.ijstr.org](http://www.ijstr.org) (Accessed: 16 August 2024).
- [21] Mayastinasari, Vita and Benyamin Lufpi (2022) ‘Efektivitas *Electronic Traffic Law Enforcement*’, *Jurnal Ilmu Kepolisian*, 16(1), pp. 63. doi:10.35879/jik.v16i1.350.
- [22] Munawar, Z. *et al.* (2022) ‘Analisis Sentimen Covid-19 Pada Media Sosial Dengan Model Neural Machine Translation’, *TEMATIK: Jurnal Teknologi Informasi Komunikasi*, 9(1), pp. 16. doi:10.38204/tematik.v9i1.917.
- [23] Agustina, Rizki (2022) *Implementasi Pemanfaatan Media Sosial Instagram Untuk Meningkatkan Prestasi Belajar Siswa Kelas Ix Pada Mata Pelajaran Pendidikan Agama Islam Di SMPN 54 Palembang*. Undergraduate Thesis, Available at: <https://repository.radenfatah.ac.id/20170/> (Accessed: 21 August 2024).
- [24] Amalia, Anugrah Resky (2021) *Studi Pengaruh Kepadatan Lalu Lintas Terhadap Pola Siklus Mengemudi Kendaraan Ringan Pada Ruas Jalan*

- Nasional Di Kota Makassar*. Thesis, Available at: <https://repository.unhas.ac.id/id/eprint/2841/> (Accessed: 21 August 2024).
- [25] Sari, R. *et al.* (2023) ‘Analisis Faktor yang Mempengaruhi Kemacetan Lalu Lintas dan Upaya Penanggulangan pada Simpang Dakota Dr. Wahidin Mataram’, *Empiricism Journal*, 4(1), pp. 242-242. doi: 10.36312/ej.v4i1.1282.
- [26] BPK (no date) *Perwali kota banjarmasin no. 23 Tahun 2016, Database Peraturan / JDIH BPK*. Available at: <https://peraturan.bpk.go.id/Details/43599> (Accessed: 22 August 2024).
- [27] Saktiyono, Heri (2021) *Pengaruh Diterapkannya Kebijakan Penutupan Jalan Protokol Saat Masa Pandemi Covid-19 Terhadap Angka Kecelakaan Lalu Lintas Di Kota Magelang*. Undergraduate Thesis. Available at: <http://eprintslib.ummgl.ac.id/3439/> (Accessed: 22 August 2024).
- [28] Yuan Sari, I. (2018) *Implementasi Kebijakan Tentang Larangan Parkir Di Badan Jalan Oleh Dinas Perhubungan Kota Palembang*. Undergraduate Thesis. Available at: <https://repository.unsri.ac.id/11513/> (Accessed: 21 August 2024).
- [29] Rufaida Hanum, R. (2019) *Merepresentasikan Makna Kata Untuk Metode Klasifikasi Naïve Bayes, Random Forest, Dan Support Vector Machine Dalam Studi Kasus Kemacetan Di Surabaya*. Undergraduate Thesis. Available at: <http://repository.its.ac.id/id/eprint/60445> (Accessed: 21 August 2024).
- [30] Anwar, Khairul (2020) *Implementasi Algoritma Naïve Bayes Untuk Memprediksi Kelayakan Kredit Nasabah*. Undergraduate Thesis. Available at: <https://repository.uinjkt.ac.id/dspace/handle/123456789/56396> (Accessed: 21 August 2024).
- [31] Fadhlullah, N. *et al.* (2022) “Penerapan teknologi web *scraping* sebagai pengumpulan data covid-19 di provinsi lampung”, *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 3(1), pp. 2. Available at: <http://jim.teknokrat.ac.id/index.php/informatika> (Accessed: 21 August 2024).
- [32] Hermanto, D.T. Setyanto, A. and Luthfi, E.T. (2021) ‘Algoritma LSTM-CNN untuk binary Klasifikasi Dengan Word2vec Pada media online’,



- Creative Information Technology Journal*, 8(1), pp. 64-64. doi:10.24076/citec.2021v8i1.264.
- [33] Qamal, M., Fuadi, W. and Muslia (2021) ‘Analisis Sentimen Toko Online Menggunakan Algoritma *Naive Bayes Classifier*’, *TTS 4.0 Jurnal Teknologi Terapan & Sains*, 2(3), pp. 22–24. Available at: <https://ojs.unimal.ac.id/tts/article/view/6771> (Accessed: 24 August 2024).
- [34] Paisal and Naraloka, T. (2021) “Pemanfaatan Teknologi Untuk Menganalisa Sentimen Masyarakat dalam Membantu Peningkatan Ekonomi Kreatif Di Era New Normal”, *Seminar Nasional Ilmu Komputer (SNASIKOM)*, 1(1), pp. 122–123. Available at: <https://proceeding.unived.ac.id/index.php/snasikom/article/view/58> (Accessed: 24 August 2024).
- [35] Rakhmanov, O. (2020) “A Comparative Study on Vectorization and Classification Techniques in Sentiment Analysis to Classify Student-Lecturer Comments”, *Procedia Computer Science*, 178, pp. 194–204. doi: 10.1016/j.procs.2020.11.021.
- [36] N. Alvi Hasanah, Nanik Suciati, and Diana Purwitasari (2021) “Pemantauan Perhatian Publik terhadap Pandemi COVID-19 melalui Klasifikasi Teks dengan Deep Learning”, *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(1), pp. 193–202. doi: 10.29207/resti.v5i1.2927.
- [37] W. Bourequat and H. Mourad (2021) “Sentiment Analysis Approach for Analyzing iPhone Release using Support Vector Machine”, *Int. J. Adv. Data Inf. Syst.*, 2(1), pp. 36–44. doi: 10.25008/ijadis.v2i1.1216.
- [38] G. K. Armah, G. Luo, and K. Qin (2014) “A Deep Analysis of the *Precision* Formula for Imbalanced Class Distribution”, *Int. J. Mach. Learn. Comput.*, 4(5), pp. 417–422. doi: 10.7763/ijmlc.2014.v4.447.
- [39] Maria Navin JR and Pankaja R. (2016) “Performance Analysis of *Text Classification* Algorithm using *Confusion Matrix*”, *Int. J. Eng. Tech. Res.*, 6(4), pp. 75–78. Available at: [https://www.erppublication.org/published\\_paper/IJETR042741.pdf](https://www.erppublication.org/published_paper/IJETR042741.pdf) (Accessed: 27 August 2024).

- [40] M. Hossin and M. Sulaiman (2015) “a Review on Evaluation *Metrics* for Data”, *Int. J. Data Min. Knowl. Manag. Process*, 5(2), pp. 1–11. doi: 10.5121/ijdkp.2015.5201.
- [41] Indonesia, P.P. (2021) Peraturan Pemerintah (PP) Nomor 30 Tahun 2021 tentang Penyelenggaraan Bidang Lalu Lintas dan Angkutan Jalan, Database Peraturan | JDIH BPK. Available at: <https://peraturan.bpk.go.id/Details/161874/pp-no-30-tahun-2021> (Accessed: 27 August 2024).
- [42] Zauhair A, M. S. (2020) *Penentuan Kondisi Kepadatan Lalu Lintas Berdasarkan Data Media Sosial Menggunakan Algoritma Naive Bayes Dibandingkan Dengan Data Hasil Rekaman Cctv Di Jalan Protokol Di Kota Palembang*. Undergraduate Thesis. Available at: <https://repository.unsri.ac.id/155089/> (Accessed: 27 August 2024).