

**Analisis Sentimen Masyarakat Indonesia Terhadap Mobil Buatan
Negara Tiongkok Pada Aplikasi X Menggunakan Metode K-
Nearest Neighbor**

SKRIPSI

Program Studi Sistem Informasi

Jenjang Sarjana



Oleh

Muhammad Fajrul Assiddiq

09031382126133

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

2025

HALAMAN PENGESAHAN

SKRIPSI

Analisis Sentimen Masyarakat Indonesia Terhadap Mobil Buatan Negara Tiongkok Pada Aplikasi X Menggunakan Metode K-Nearest Neighbor

Sebagai salah satu syarat untuk penyelesaian studi di

Program Studi S1 Sistem Informasi

Oleh:

MUHAMMAD FAJRUL ASSIDDIQ

09031382126133

Pembimbing 1

: Dr. Ali Ibrahim, S.Kom., M.T.

NIP. 198407212019031004

Mengetahui

Ketua Jurusan Sistem Informasi



Ahmad Rifai, S.T., M.T.

1979102010121003

HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan di bawah ini:

Nama : Muhammad Fajrul Assiddiq

NIM : 09031382126133

Program Studi : Sistem Informasi Bilingual

Judul : Analisis Sentimen Masyarakat Indonesia Terhadap Mobil Buatan
Negara Tiongkok Pada Aplikasi X Menggunakan Metode K-
Nearest Neighbor

Hasil Pengecekan *Software iThenticate/Turnitin*: 11%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya saya sendiri dan bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku. Demikianlah pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 23 Mei 2025

Penulis,



Muhammad Fajrul Assiddiq

NIM. 09031382126133

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Jumat
Tanggal : 23 Mei 2025
Nama : Muhammad Fajrul Assiddiq
NIM : 09031382126133
Judul Tugas Akhir : Analisis Sentimen Masyarakat Indonesia Terhadap Mobil
Buatan Negara Tiongkok Pada Aplikasi X Menggunakan
Metode *K-Nearest Neighbor*

Komisi Penguji :

1. Ketua : Dr. Fathoni, MMSI.
2. Penguji : Mira Afrina, M. Sc., Ph.D
3. Pembimbing : Dr. Ali Ibrahim, S.Kom., M.T.



Mengetahui,
Ketua Jurusan Sistem Informasi



Ahmad Rifai, S.T., M.T.
NIP. 197910202010121003

HALAMAN PERSEMBAHAN

MOTTO

"Fortune Favors the Brave"

– Pliny the Elder

“Keberhasilan tidak selalu datang dari mereka yang paling pintar atau paling cepat, tetapi dari mereka yang berani melangkah dan tidak berhenti. Keberanian untuk mencoba adalah langkah pertama menuju keberhasilan.”

Skripsi ini dipersembahkan untuk:

- ❖ Orang Tua dan Keluarga
- ❖ Diri Sendiri
- ❖ Dosen Pembimbing
- ❖ Rekan Seperjuangan
- ❖ Universitas Sriwijaya

KATA PENGANTAR

Alhamdulillah, segala puji dan syukur kepada Allah yang Maha Esa, atas berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul **“Analisis Sentimen Masyarakat Indonesia Terhadap Mobil Buatan Negara Tiongkok pada Aplikasi X Menggunakan Metode K-Nearest Neighbor”** guna memenuhi salah satu persyaratan dalam menyelesaikan pendidikan pada Jurusan Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya.

Di kesempatan ini penulis juga ingin menyampaikan ucapan terima kasih yang sebesar besarnya kepada para pihak yang telah memberikan dukungan dalam menyelesaikan tugas akhir ini baik secara langsung maupun secara tidak langsung,. Penulis ingin menyampaikan rasa terima kasih kepada:

1. Allah Subhanahu Wa Ta’ala Yang Maha Pengasih dan Maha Penyayang, yang telah mencurahkan berkah dan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi ini.
2. Kedua orang tua saya dan adik-adik saya tercinta sebagai penyemangat utama di balik layar. Terima kasih atas segala doa yang tak pernah putus, semangat yang selalu kalian kirimkan.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Ahmad Rifai, M.T selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Bapak Dr. Ali Ibrahim, S.Kom., M.T., selaku dosen pembimbing tugas akhir yang dengan sabar membimbing, mengarahkan, dan bahkan rela ditemui

kapan saja dan selalu memudahkan perjalanan pengerjaan skripsi saya. Terima kasih atas kesabaran Bapak menghadapi saya yang kadang suka menghilang dalam waktu yang lama.

6. Seluruh Bapak dan Ibu Dosen Fakultas Ilmu Komputer Universitas Sriwijaya khususnya Dosen Jurusan Sistem Informasi yang telah membimbing dan memberikan ilmu pengetahuan dan pengalaman selama masa studi.
7. Ibu Rifka selaku Admin Jurusan Sistem Informasi Bilingual yang telah membantu proses administrasi dan memberikan informasi hingga akhir perkuliahan.
8. Rekan-rekan kuliah seperjuangan, khususnya rekan kelas Sistem Bilingual A 2021, yang telah menjadi teman diskusi, belajar, saling mendukung, dan berbagi cerita suka maupun duka selama masa perkuliahan. Terima kasih atas kebersamaan dan semangat yang telah menjadi bagian penting dalam perjalanan akademik penulis.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kata sempurna, maka dari itu, penulis dengan tulus membuka diri untuk menerima kritik dan saran yang membangun demi kesempurnaan penulisan tugas akhir pada masa yang akan datang. Semoga Tugas Akhir ini bermanfaat dalam menambah wawasan serta menjadi acuan untuk penelitian serupa nantinya.

Palembang, 23 Mei 2025
Penulis,

Muhammad Fajrul Assiddiq
NIM. 09031382126133

**Analisis Sentimen Masyarakat Indonesia Terhadap Mobil Buatan Negara
Tiongkok Pada Aplikasi X Menggunakan Metode K-Nearest Neighbor**

Oleh

Muhammad Fajrul Assiddiq

090313821216133

ABSTRAK

Citra barang buatan Tiongkok sejak lama dikenal negatif di masyarakat, terutama terkait kualitas produk yang dianggap mudah rusak dan kurang tahan lama. Namun, dalam beberapa tahun terakhir, produk otomotif asal Tiongkok mulai memasuki pasar Indonesia dengan inovasi dan harga yang kompetitif, khususnya di segmen kendaraan listrik. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat Indonesia terhadap mobil buatan Tiongkok yang diungkapkan melalui media sosial X dengan menggunakan metode K-Nearest Neighbor (KNN). Data yang digunakan berupa tweet berbahasa Indonesia selama tiga tahun terakhir yang telah melalui proses pre-processing meliputi pembersihan data, normalisasi, case folding, tokenisasi, dan penghapusan stopword. Labeling sentimen dilakukan secara otomatis menggunakan model pre-trained IndoBERT, lalu fitur teks diekstraksi dengan TF-IDF. Algoritma KNN diaplikasikan untuk mengklasifikasikan sentimen menjadi positif dan negatif dengan parameter yang telah dioptimalkan melalui hyperparameter tuning. Hasil analisis menunjukkan bahwa mayoritas masyarakat Indonesia memberikan sentimen positif sebesar 93,49% terhadap mobil Tiongkok. Model KNN mencapai akurasi sebesar 86%, meskipun performa dalam mendeteksi sentimen negatif masih perlu ditingkatkan. Sentimen positif didominasi oleh faktor harga yang terjangkau dan inovasi teknologi kendaraan listrik, sementara sentimen negatif berkaitan dengan kekhawatiran terhadap kualitas produk dan layanan purna jual. Penelitian ini memberikan kontribusi praktis dalam memahami persepsi konsumen dan menjadi dasar bagi pelaku industri otomotif untuk memperbaiki strategi pemasaran dan membangun citra merek yang lebih baik.

Kata Kunci: Analisis Sentimen, K-Nearest Neighbor, Mobil Tiongkok, Media Sosial X, IndoBERT, TF-IDF.

Sentiment Analysis of Indonesian Public Toward Chinese-Made Cars on X Application Using the K-Nearest Neighbor Method

By

Muhammad Fajrul Assiddiq

09031382126133

ABSTRACT

The image of Chinese-made goods has long been perceived negatively by the public, especially regarding product quality, which is often considered fragile and not durable. However, in recent years, Chinese automotive products have begun to enter the Indonesian market with innovations and competitive pricing, particularly in the electric vehicle segment. This study aims to analyze Indonesian public sentiment toward Chinese-made cars expressed through the social media platform X using the K-Nearest Neighbor (KNN) method. The data were obtained from Indonesian-language tweets over the past three years and processed through several pre-processing stages, including cleaning, normalization, case folding, tokenization, and stopword removal. Sentiment analysis was conducted using the IndoBERT pre-trained model, followed by feature extraction using TF-IDF. The KNN algorithm was applied to classify sentiments into positive or negative classes by tuning its parameters using hyperparameter tuning techniques. The results showed that most Indonesians held negative sentiments, with 93.49% of tweets expressing unfavorable opinions toward Chinese-made cars. The KNN model achieved an accuracy of 86%, despite its challenges in detecting subtle negative sentiments. Positive sentiment was generally associated with low prices and technological innovation in electric vehicles, while negative sentiment focused on durability issues and poor service quality. This study contributes practical insights into consumer sentiment, which can be used by automotive industry stakeholders to improve strategies and enhance the image of foreign-made vehicles in the Indonesian market.

Keywords: Sentiment Analysis, K-Nearest Neighbor, Chinese Cars, X Social Media, IndoBERT, TF-IDF.

DAFTAR PUSTAKA

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN BEBAS PLAGIAT.....	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR PUSTAKA.....	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	8
2.1 Penelitian Terdahulu.....	8
2.2 Analisis Sentimen.....	15
2.3 Industri Otomotif Tiongkok.....	17
2.4 Media Sosial <i>X</i>	18
2.5 <i>Natural Language Processing</i>	19
2.6 <i>Text Mining</i>	20
2.7 <i>Labeling</i>	20
2.8 <i>Feature Extraction</i>	21
2.9 Metode K-Nearest Neighbor (KNN)	22
2.10 <i>Confusion Matrix</i>	23
BAB III METODOLOGI PENELITIAN	25
3.1 Kerangka Penelitian	25
3.2 Identifikasi Masalah	25
3.3 Studi Literatur	26
3.4 <i>Data Crawling</i>	26
3.5 <i>Pre-processing Data</i>	26

3.6 <i>Data Splitting</i>	27
3.7 <i>Labeling</i>	27
3.8 <i>Feature Extraction</i>	28
3.9 Algoritma <i>K-Nearest Neighbor</i> (KNN)	29
3.10 Evaluasi.....	29
BAB IV HASIL DAN PEMBAHASAN.....	30
4.1 <i>Crawling Data</i>	30
4.2 Pre-Processing Data	33
4.2.1 <i>Text Cleaning</i>	33
4.2.2 <i>Normalization</i>	35
4.2.3 <i>Case Folding</i>	38
4.2.4 <i>Tokenization</i>	41
4.2.5 <i>Stopword Removal</i>	44
4.3 <i>Data Splitting</i>	47
4.4 <i>Labeling dengan IndoBERT</i>	48
4.5 <i>Ekstraksi Fitur dengan TF-IDF</i>	51
4.6 <i>Penerapan Algoritma K-Nearest Neighbors (KNN)</i>	52
4.7 Evaluasi.....	53
BAB V KESIMPULAN DAN SARAN	60
5.1 <i>Kesimpulan</i>	60
5.2 <i>Saran</i>	61
DAFTAR PUSTAKA.....	63
L A M P I R A N	67

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Analisis Sentimen.....	15
Gambar 3. 1 Tahapan Penelitian.....	25
Gambar 4. 1 Contoh Sentimen Pada Media Sosial X.....	31
Gambar 4. 2 Jumlah Data Crawl Tweet.....	31
Gambar 4. 3 Jumlah Data Tweet yang Akan Digunakan.....	32
Gambar 4. 4 Proses Data Cleaning.....	34
Gambar 4. 5 Proses <i>Normalization</i>	36
Gambar 4. 6 Proses Case FoldingTabel 4.4 Contoh Data Setelah Proses Case Folding	39
Gambar 4. 7 Proses <i>Tokenization</i>	43
Gambar 4. 8 Proses Stopword Removal.....	46
Gambar 4. 9 Tampilan Data Latih	48
Gambar 4. 10 Tampilan Data Uji.....	48
Gambar 4. 11 Tampilan Data Latih Setelah Proses Labeling dengan Model <i>Pre- trained</i> IndoBert.....	50
Gambar 4. 12 Tampilan Distribusi Sentimen	51
Gambar 4. 13 Hasil Klasifikasi Label Sentimen model KNN.....	53
Gambar 4. 14 Distribusi Sentimen Positif dan Negatif	53
Gambar 4. 15 Tampilan Metrik Evaluasi Model	55
Gambar 4. 16 Tampilan Word Cloud Label Positif	58
Gambar 4. 17 Tampilan Word Cloud Label Negatif.....	59

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix.....	23
Tabel 4. 1 Contoh Data Tweet setelah Proses Crawling.....	32
Tabel 4. 2 Contoh Data setelah Proses <i>Cleaning</i>	34
Tabel 4. 3 Contoh Data Setelah Proses <i>Normalization</i>	36
Tabel 4. 4 Contoh Data Setelah Proses <i>Tokenization</i>	43
Tabel 4. 5 Contoh Data Setelah Proses <i>Stopword Removal</i>	46
Tabel 4. 6 Tabel Confusion Matrix	56

DAFTAR LAMPIRAN

Lampiran 1 Dataset Hasil Sentimen.....	A-1
Lampiran 2 Surat Kesediaan Dosen Pembimbing.....	B-1
Lampiran 3 Surat Keputusan Tugas Akhir	C-1
Lampiran 4 <i>Similarity Check</i> Skripsi	D-1
Lampiran 5 Surat Keterangan Pengecekan <i>Similarity</i> Skripsi	E-1
Lampiran 6 Formulir Perbaikan Komprehensif	F-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Industrialisasi otomotif Indonesia telah mengalami pertumbuhan pesat dalam beberapa dekade terakhir. Menurut data dari Gabungan Industri Kendaraan Bermotor (GAIKINDO), penjualan mobil secara *wholesales* (dari pabrik ke dealer) tercatat sebesar 865.723 pada tahun 2024 (gaikindo.or.id, 2025). Berbagai merek mobil dari beragam negara berlomba-lomba untuk memasuki pasar Indonesia yang besar dan terus berkembang. Salah satu fenomena yang menarik adalah banyaknya merek-merek mobil merek Tiongkok yang masuk ke Indonesia. Sejak tahun 2023, tercatat enam merek mobil baru asal Tiongkok yang memasuki pasar Indonesia, yaitu Neta, Haval, Tank, Ora, Maxus, dan juga BYD (Nanda & Ravel Stanly, 2023). Pada tahun 2025, merek BYD, Chery, Wuling, dan Denza mendapatkan angka penjualan unit yang relatif tinggi pada periode April 2025 (inews.id, 2025). Kehadiran mobil asal Tiongkok di Indonesia dipengaruhi oleh beberapa faktor yaitu, harga yang ditawarkan lebih terjangkau dengan fitur-fitur inovatif, serta membuat diversifikasi produk dengan memperkuat posisi di segmen mobil listrik sehingga menarik bagi konsumen Indonesia (Satria & Maulana, 2024).

Meskipun demikian, persepsi masyarakat Indonesia terhadap merek mobil Tiongkok masih beragam. Beberapa kalangan menganggap mobil buatan Tiongkok memiliki harga yang lebih murah dibandingkan kompetitor dengan fitur yang cukup menarik. Namun, ada pula yang meragukan kualitas, dan ketahanan dari produk tersebut. Perbedaan persepsi inilah yang menciptakan sentimen positif, negatif, maupun netral di kalangan konsumen.

Di era digital saat ini, sentimen masyarakat banyak diekspresikan melalui media sosial, forum online, maupun platform e-commerce. Data-data tersebut merupakan sumber informasi yang kaya untuk diteliti, terutama untuk mendeteksi opini, emosi, dan persepsi publik terhadap merek atau produk tertentu. Dalam konteks ini, Analisis sentimen sangat berguna untuk menganalisis opini publik pengguna media sosial X tentang berbagai isu atau topik yang sedang ramai dibicarakan (Permana et al., 2022).

Analisis sentimen merupakan proses untuk mengidentifikasi perasaan atau emosi seseorang terhadap suatu objek, yang kemudian menghasilkan sentimen negatif, positif, atau netral (Budiman et al., 2024). Analisis sentimen mengacu pada proses pengelompokan ulasan atau teks berdasarkan arah sentimen, baik positif, netral, maupun negatif. Secara teknis, analisis sentimen melibatkan proses pengolahan data secara terstruktur untuk mengidentifikasi, mengekstraksi, dan mengukur keadaan afektif dari suatu objek atau target. Teknik yang digunakan dapat bervariasi tergantung pada jenis data dan tujuan analisis, namun umumnya mencakup tahap penambangan data (*data mining*), pembersihan data (*data cleaning*), pengolahan data (*data processing*), dan penyajian hasil (*data presentation*). Teknik yang digunakan bervariasi berdasarkan data dan tujuan analisis, namun umumnya melibatkan fase penambangan data (*data mining*), pembersihan data (*data cleaning*), pengolahan data (*data processing*), dan penyajian hasil (*data presentation*). Salah satu pendekatan yang dapat diterapkan adalah analisis sentimen menggunakan teknik *text mining*. *Text mining* merupakan proses teknik komputasi dan pembelajaran mesin yang digunakan untuk mengolah dan menganalisis teks secara otomatis (Amna et al., 2023). Secara teknis, *text*

mining adalah proses ekstraksi informasi dari data berupa teks. Sumber data biasanya didapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen (Munthe et al., 2022)

Solusi algoritma yang penulis pilih adalah K-Nearest Neighbor. Penulis memilih algoritma ini karena sebelumnya sudah melakukan tinjauan pustaka dari metode yang serupa, di antaranya sebagai berikut. Penelitian sebelumnya yang dilakukan oleh (Mara et al., 2021) menunjukkan Kelebihan algoritma K-NN terletak pada kemampuannya untuk mencapai akurasi tinggi, yang terbukti efektif dalam berbagai aplikasi. Penelitian ini bertujuan untuk mengklasifikasikan penilaian mahasiswa terhadap pandemi, baik berupa opini negatif maupun positif, serta menghitung akurasi algoritma modeling K-NN. Hasil penelitian menunjukkan akurasi yang didapat sebesar 87% dan untuk AUC sebesar 91,6%. Dalam penelitian yang dikerjakan oleh Sepriadi et al. (2023), metode KNN diterapkan dalam pengelompokkan opini masyarakat terhadap aplikasi MyPertamina. Perbandingan pemisahan data latih dan data uji adalah 80:20 dan nilai $K=9$, akurasi mencapai 73%, presisi 70%, dan *recall* 73%. Hasilnya menunjukkan bahwa akurasi tertinggi diperoleh dengan pembagian porsi data 80:20 dengan nilai $K=9$, menghasilkan nilai akurasi 73%, presisi 70%, dan *recall* 73% tanpa menggunakan *undersampling*. Dengan menggunakan *undersampling* memakai perbandingan data yang sama, akurasi menurun menjadi 57%, presisi meningkat menjadi 73%, namun *recall* menurun ke 57%. Penurunan akurasi ini disebabkan oleh berkurangnya jumlah data latih akibat penerapan *undersampling*. Kelebihan K-NN terletak pada prinsipnya yang sederhana, di mana algoritma ini bekerja dengan membandingkan jarak

terpendek antara sampel uji dan sampel latih. Dan penelitian yang dilakukan oleh Pamungkas & Kharisudin (2021) membandingkan tiga algoritma klasifikasi, yaitu Support Vector Machine (SVM), Naive Bayes, dan K-Nearest Neighbor (K-NN), dalam menganalisis sentimen review produk. Hasil penelitian menunjukkan bahwa algoritma K-NN memiliki keunggulan dalam hal akurasi dibandingkan dengan dua algoritma lainnya, terutama pada dataset dengan distribusi kelas yang seimbang.

Berdasarkan uraian yang telah dipaparkan. Penulis memilih metode K-Nearest Neighbor karena metode ini memiliki beberapa keunggulan dalam analisis sentimen, di antaranya adalah akurasi yang tinggi. K-NN juga tahan terhadap data *noise* dan efektif untuk dataset berukuran besar, karena bekerja dengan prinsip perhitungan jarak antar sampel. Selain itu, metode ini memberikan hasil klasifikasi yang baik dengan pemilihan parameter k yang tepat. Penelitian juga menunjukkan bahwa K-NN menunjukkan kinerja yang lebih unggul dibandingkan dengan beberapa metode lain, seperti Naïve Bayes dan Support Vector Machine (SVM). Dengan teknik optimasi seperti *cosine similarity* dan pemrosesan teks yang baik, K-NN dapat menghasilkan prediksi sentimen yang akurat dan stabil dalam berbagai domain, termasuk ulasan produk, opini publik di media sosial, dan analisis terhadap berbagai isu sosial.

Atas dasar penjelasan di atas, penulis melakukan penelitian yang berjudul **“ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP MOBIL BUATAN TIONGKOK PADA APLIKASI X MENGGUNAKAN METODE K-NEAREST NEIGHBOR”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang penulis jabarkan di atas, rumusan masalah yang dapat diidentifikasi adalah sebagai berikut:

1. Bagaimana sentimen masyarakat Indonesia (positif atau negatif) terhadap mobil merek Tiongkok berdasarkan data yang diperoleh dari X?
2. Faktor-faktor apa saja yang memengaruhi sentimen masyarakat tersebut terhadap mobil merek Tiongkok?
3. Seberapa efektif metode K-Nearest Neighbor dalam menganalisis sentimen masyarakat Indonesia terhadap mobil merek Tiongkok?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan di atas, penelitian ini bertujuan untuk:

1. Mengklasifikasikan sentimen masyarakat Indonesia terhadap mobil merek Tiongkok menjadi *review* sentimen negatif dan positif menggunakan metode K-Nearest Neighbor.
2. Mengetahui faktor-faktor yang memengaruhi sentimen masyarakat tersebut.
3. Mengevaluasi kinerja prediksi K-Nearest Neighbor dalam menganalisis data sentimen di aplikasi X.

1.4 Manfaat Penelitian

Penulis berharap penelitian ini dapat memberikan manfaat sebagai berikut:

1. Menambah wawasan dalam bidang analisis sentimen, khususnya dengan penerapan metode K-Nearest Neighbor.

2. Memberikan referensi bagi mahasiswa atau peneliti lain yang ingin mengkaji metode-metode *data mining* khususnya metode K-Nearest Neighbor dalam analisis sentimen.
3. Memberikan informasi penting bagi pelaku industri otomotif untuk merancang strategi pemasaran yang lebih efektif berdasarkan sentimen masyarakat.
4. Membantu masyarakat memahami persepsi umum terkait mobil buatan asal Tiongkok melalui data sentimen yang telah diolah agar menjadi bahan pertimbangan saat ingin mengambil keputusan pembelian.

1.5 Batasan Masalah

Untuk memastikan fokus dan kejelasan pada penelitian ini, berikut batasan masalah yang penulis tetapkan:

1. Data yang digunakan dalam penelitian ini hanya mencakup *tweet* dan komentar dari pengguna aplikasi **X** yang relevan dengan topik mobil buatan asal Tiongkok.
2. Data dikumpulkan dalam kurun 3 tahun terakhir untuk menjaga relevansi.
3. Analisis sentimen dilakukan hanya dengan menggunakan metode K-Nearest Neighbor untuk klasifikasi sentimen.
4. Penelitian terbatas pada konteks masyarakat Indonesia dan tidak membahas persepsi masyarakat di negara lain terhadap mobil buatan asal Tiongkok.
5. Analisis difokuskan pada sentimen positif dan negatif tanpa mendalami emosi spesifik seperti marah, bahagia, atau kecewa.
6. Penelitian bergantung pada data yang tersedia di aplikasi **X** dan tidak mencakup data dari media sosial lain, seperti Tiktok atau Instagram.

7. Fokus hanya pada mobil merk Tiongkok dan tidak mencakup perbandingan langsung dengan mobil dari negara lain.

DAFTAR PUSTAKA

- Abimanyu, D., Budianita, E., Pandu Cynthia, E., Yanto, F., Studi Teknik Informatika, P., & Sains Dan Teknologi, F. (2022). Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER. *Jurnal Nasional Komputasi Dan Teknologi Informasi*, 5(3). <https://techno.kompas.com>
- Agung, S. (2024). *Implementasi Text Mining untuk Analisis Review pada Aplikasi Crowdfunding LX dan ST Menggunakan Metode Sentiment Analysis*. 2(1), 124–130. <https://doi.org/10.35870/ljit.v2i1.2245>
- Amna, Wahyuddin S, Sudipa, I. G. I., Tri, A. E., & Wahidin Putra Ahmad Jurnaidi. (2023). Text Mining. In *Data Mining*. PT GLOBAL EKSEKUTIF TEKNOLOGI.
- antaranews.com. (2024, February 20). *Beijing klaim produk otomotif China murah karena produksi efisien*. Antaranews.Com. <https://www.antaranews.com/berita/3974295/beijing-klaim-produk-otomotif-china-murah-karena-produksi-efisien>
- Apriliani, D., Susanto, A., Fikri Hidayattullah, M., & Wiro Sasmito, G. (2023). *Sentimen Analisis Pandangan Masyarakat Terhadap Vaksinasi Covid 19 Menggunakan K-Nearest Neighbors*. 8(1).
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *FACCT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Budiman, B., Anggraeni, Z., Habibi, C., & Alamsyah, N. (2024). Analisis Sentimen Publik pada Media Sosial Twitter Terhadap Tiket.com Menggunakan Algoritma Klasifikasi. *Jurnal Informatika*, 11, 1–10. <https://doi.org/10.31294/inf.v11i1.17988>
- Dani, Y., & Ginting, M. A. (2023). *Classification of Predicting Customer Ad Clicks Using Logistic Regression and k-Nearest Neighbors*. www.joiv.org/index.php/joiv
- Firdaus, R. Z., Wijoyo, S. H., & Purnomo, W. (2025). Analisis Sentimen Berbasis Aspek Ulasan Pengguna Aplikasi Alfagift Menggunakan Metode Random Forest dan Pemodelan Topik Latent Dirichlet Allocation. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 9. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/14482/6461>
- Fudholi, L. A., Rahaningsih, N., & Dana, R. D. (2024). SENTIMEN ANALISIS PERILAKU PENGEMAR COLDPLAY DI MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAIVE BAYES. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 3).
- gaikindo.or.id. (2025, January). *Proyeksi GAIKINDO untuk Penjualan Mobil Indonesia 2025*. Gaikindo.or.Id. <https://www.gaikindo.or.id/proyeksi-gaikindo-untuk-penjualan-mobil-indonesia-2025/>

- 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).
- Habib Kusuma, I., & Cahyono, N. (2023). *Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor*. 8(3).
- Hasri, C. F., & Alita, D. (2022). PENERAPAN METODE NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TERHADAP DAMPAK VIRUS CORONA DI TWITTER. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, 3(2), 145–160. <http://jim.teknokrat.ac.id/index.php/informatika>
- Hazimi, M. S. (2023). *Analisis Sentimen Masyarakat Terhadap Calon Kandidat Presiden dalam Persiapan Pemilu Tahun 2024 dengan Menggunakan Metode K-Nearest Neighbor dan Opstipasi Particle Swarm Optimization*.
- imd.org. (2024, January 26). *China's automotive odyssey: From joint ventures to global EV dominance*. Imd.Org. <https://www.imd.org/ibyimd/innovation/chinas-automotive-odyssey-from-joint-ventures-to-global-ev-dominance/>
- inews.id. (2025, May 12). *Daftar Merek Mobil China Terlaris di Indonesia April 2025, Ini 10 Brand Paling Digemari*. https://www.inews.id/otomotif/motor/daftar-merek-mobil-china-terlaris-di-indonesia-april-2025-ini-10-brand-paling-digemari?utm_source=chatgpt.com
- Khoirul, M., Hayati, U., & Nurdiawan, O. (2023). ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 1).
- Mara, A. A. P. T., Sedyono, E., & Purnomo, H. (2021). Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Metode Pembelajaran Dalam Jaringan (DARING) Di Universitas Kristen Wira Wacana Sumba. In *JOINTER-JOURNAL OF INFORMATICS ENGINEERING* (Vol. 01, Issue 02).
- Merdiansah, R., & Ali Ridha, A. (2024). Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 7(1), 221–228.
- Munthe, C., Hasibuan, N., & Hutabarat, H. (2022). Penerapan Algoritma Text Mining Dan TF-RF Dalam Menentukan Promo Produk Pada Marketplace. *Resolusi : Rekayasa Teknik Informatika Dan Informasi*, 2, 110–115. <https://doi.org/10.30865/resolusi.v2i3.309>
- Nanda, A. M., & Ravel Stanly. (2023, December 19). *Lima Merek Mobil Baru yang Masuk Indonesia Sepanjang 2023*. Kompas.Com. <https://otomotif.kompas.com/read/2023/12/19/144200515/lima-merek-mobil-baru-yang-masuk-indonesia-sepanjang-2023?>
- Pamungkas, F. S., & Kharisudin, I. (2021). *Analisis Sentimen dengan SVM*. 4, 628–634. <https://journal.unnes.ac.id/sju/index.php/prisma/>

- Permana, A. A., Fahrezi, M. F., Priyanggodo, D. Y., Kristiyanti, D. A., Sihotang, M., Studi, P., Informatika, T., Teknik, F., Muhammadiyah, U., Jalan, T., Kemerdekaan, P., & Tangerang, C. K. (2022). *SENTIMEN ANALISIS OPINI MASYARAKAT PADA MEDIA SOSIAL TWITTER TERHADAP VAKSIN BERBAYAR MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER (NBC) SENTIMENT ANALYSIS OF PUBLIC OPINION ON SOCIAL MEDIA TWITTER ON PAID VACCINE USING NAÏVE BAYES CLASSIFIER (NBC) METHOD*. 10(2), 84–92. <http://jurnal.umt.ac.id/index.php/jt/index>
- Putra, G. G. S., Swastika, W., & Irawan, P. L. T. I. (2022). Perbandingan Particle Swarm Optimization dengan Genetic Algorithm dalam Feature Selection untuk Analisis Sentimen pada Permendikbudristek PPKS-LPT. *Jurnal Edukasi Dan Penelitian Informatika*.
- Ramadhani, B., & Suryono, R. R. (2024). Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(2), 714. <https://doi.org/10.30865/mib.v8i2.7458>
- Ridwansyah, T. (2022). Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan Naïve Bayes Classifier. *Media Online*, 2(5), 178–185. <https://djournals.com/klik>
- Safira, A., & Hasan, F. N. (2023). ANALISIS SENTIMEN MASYARAKAT TERHADAP PAYLATER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER. *Jurnal Sistem Informasi*, 5(1).
- Safitri, T., Umaidah, Y., & Maulana, I. (2023). Analisis Sentimen Pengguna Twitter Terhadap BTS Menggunakan Algoritma Support Vector Machine. In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 7, Issue 1). <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Said, R., & Faraby, A. (2023). *Perbandingan Algoritma Machine Learning untuk Analisis Sentimen Berbasis Aspek pada Review Female Daily* (Vol. 10, Issue 3).
- Salsabila, N. A. (2022). *ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP TOKOH GUS DUR MENGGUNAKAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE (SVM)*.
- Saputra, H. (2023). ANALISIS SENTIMEN PADA VAKSIN BOOSTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE MULTICLASS DI TWITTER. In *Teknologipintar.org* (Vol. 3, Issue 10).
- Satria, G., & Maulana, A. (2024, November 5). *Chery Klaim Mobil China Makin Diterima di Indonesia*. Kompas.Com. <https://otomotif.kompas.com/read/2024/11/05/141200815/chery-klaim-mobil-china-makin-diterima-di-indonesia?>
- Sepriadi, N., Budianita', E., Fikry, M., Fakultas, P., Sains, ", Teknologi, D., Negeri, I., & Riau, K. (2023). *INFORMASI (Jurnal Informatika dan Sistem Informasi) Analisis Sentimen Review Aplikasi MyPertamina menggunakan Word Embedding Fasttext dan Algoritma K-Nearest Neighbor*.

Supriyanto, J., Alita, D., & Isnain, A. R. (2023). Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 4(1), 74–80.
<https://doi.org/10.33365/jatika.v4i1.2468>

Wahyuni, W. (2022). Implementation of the Support Vector Machine Method for Sentiment Analysis Using Twitter Data. *International Journal of Knowledge in Database*, 02(02), 166–180. <https://doi.org/10.30983/ijokid.v2i2.6019>

x.com. (n.d.). *About X*. X.Com. Retrieved February 5, 2025, from <https://about.x.com/en>

Zaqi Taufan, A., & Wibowo, W. (n.d.). ANALISIS SENTIMEN TERKAIT PERSEPSI KEAMANAN DATA INFORMASI DAN PRIVASI DI INDONESIA MENGGUNAKAN PENDEKATAN MACHINE LEARNING. *Jurnal Informatika Teknologi Dan Sains*.