



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



KLASIFIKASI SENTIMEN APLIKASI TOKOPEDIA DENGAN MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER BERDASARKAN ULASAN DI GOOGLE PLAY

TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

PUTRI APRIYAWATI
NIM. 12050127491



UIN SUSKA RIAU

UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS ISLAM NEGERI

SULTAN SYARIF KASIM RIAU

PEKANBARU

2025



1. Dianggap mengump sebagai atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dianggap mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

KLASIFIKASI SENTIMEN APLIKASI TOKOPEDIA DENGAN MENGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERDASARKAN ULASAN DI GOOGLE PLAY

TUGAS AKHIR

Oleh

PUTRI APRIYAWATI
12050127491

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, 06 Januari 2026

Pembimbing I

Yusra, S.T., M.T
NIP. 198401232015032001

Pembimbing II

Muhammad Fikry, S.T., M.Sc
NIP. 198010182007101002



LEMBAR PENGESAHAN

KLASIFIKASI SENTIMEN APLIKASI TOKOPEDIA DENGAN MENGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERDASARKAN ULASAN DI GOOGLE PLAY

TUGAS AKHIR

Oleh

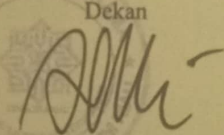
PUTRI APRIYAWATI

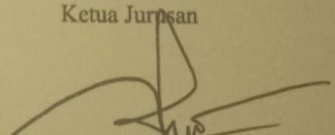
12050127491

Telah dipertahankan di depan dewan penguji
sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 6 Januari 2026

Pekanbaru, 6 Januari 2026

Mengesahkan
Ketua Jurusan


Dr. Yuslenita Muda, S.Si., M.Sc
NIP. 197701032007102001


Muhammad Amandes, S.T., M.T
NIP. 198612062015031004

DEWAN PENGUJI

Ketua	: Elvia Budianita, S.T., M.Cs
Pembimbing I	: Yusra, S.T., M.T
Pembimbing II	: Muhammad Fikry, S.T., M.Sc
Penguji I	: Rahmad Abdillah, ST., M.T
Penguji II	: Fadhilah Syafria, S.T., M.Kom

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Putri Apriyawati

NIM : 12050127491

Tempat/Tgl.Lahir : Selatpanjang, 23 April 2002

Fakultas : Sains dan Teknologi

Prodi : Teknik Informatika

Judul Skripsi : "Klasifikasi Sentimen Aplikasi Tokopedia dengan menggunakan Metode Naïve Bayes Classifier berdasarkan ulasan di Google Play"

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan Skripsi dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Skripsi saya ini, saya nyatakan bebas dari plagiat
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan Skripsi saya tersebut, maka saya bersedia menerima sanksi sesuai perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru
Yang membuat pernyataan,



Putri Apriyawati
NIM. 12050127491



LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Selama penulis memiliki hak cipta, proyek akhir yang belum dipublikasikan ini terdaftar dan dapat diakses di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim di Riau. Referensi diperbolehkan, tetapi ringkasan atau kutipan hanya dapat diberikan dengan persetujuan penulis dan harus mengikuti pedoman ilmiah saat merujuk sumber.

Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim di Riau harus memberikan otorisasi untuk setiap reproduksi atau publikasi bagian mana pun dari proyek akhir ini. Nama, nomor identifikasi peminjaman, dan tanggal peminjaman harus diberikan oleh perpustakaan yang meminjamkan proyek akhir ini kepada anggotanya.

Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Tugas Akhir ini dengan judul “JUDUL TUGAS AKHIR” adalah gagasan asli dari saya sendiri dan belum pernah dijadikan Tugas Akhir atau sejenisnya di Universitas Islam Negeri Sultan Syarif Kasim Riau maupun di perguruan tinggi lain.
2. Dalam Tugas Akhir ini TIDAK terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali tertulis dengan jelas dan dicantumkan sebagai referensi di dalam Daftar Pustaka.
3. Dalam Tugas Akhir ini TIDAK terdapat penggunaan Kecerdasan Buatan Generatif (Generative AI) yang bertentangan dengan ketentuan dan peraturan yang berlaku.
4. Saya bersedia menerima sanksi sesuai ketentuan yang berlaku apabila di kemudian hari terbukti bahwa Tugas Akhir ini melanggar kode etik maupun peraturan yang berlaku, termasuk plagiat ataupun pelanggaran hak cipta.

Demikianlah pernyataan ini dibuat dengan sebenarnya.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN

Bismillahirrahmanirrahim.

Dengan menyebut nama Allah yang Maha Pengasih lagi Maha Penyayang. Segala puji bagi Allah SWT. Atas limpahan rahmat, nikmat, kekuatan dan kesempatan yang diberikan hingga akhirnya skripsi ini dapat terselesaikan. Shalawat dan salam kepada baginda Muhammad SAW yang telah menjadi suri teladan yang baik bagi hamba. Dengan penuh syukur dan ketulusan, karya ini penulis persembahkan kepada :

1. Cinta pertama dan panutan hidupku, Bapak Harianto dan ibunda tercinta Irawati. Terimakasih atas cinta yang tiada henti dan pengorbanan yang tak terhingga. Meski beliau tak sempat merasakan bangku perkuliahan, namun beliau mampu memberikan segalanya hingga penulis mampu merasakan bangku perkuliahan dan menyelesaikan studinya sampai meraih gelar sarjana pertama dikeluarga dan keluarga besar Alm. Idrus. Terimakasih sudah mengajarkan arti perjuangan dan pengorbanan.
2. Adikku tercinta, Ressa Anggun Apriyani, terimakasih sudah ikut serta dalam proses penulis menempuh pendidikan selama ini, terimakasih atas semangat, doa dan cinta yang selalu diberikan kepada penulis. Tumbuh lebih baik dibanding diriku.
3. Keluarga besar Alm. Idrus yang dari awal perkuliahan selalu memberi semangat, dukungan dan doa. Dukungan tersebut menjadi motivasi berharga bagi penulis dalam menyelesaikan studi ini.
4. Sahabat penulis yang selalu mendengar keluh kesah saya dan selalu menjadi penyemangat ketika penulis mengalami masa sulit.
5. Terimakasih kepada bapak, ibu dosen pembimbing dan penguji yang tidak kenal lelah dalam membimbing saya menyelesaikan tugas akhir ini.
6. Untuk semua pihak yang telah membantu saya dalam penyelesaian tugas akhir ini yang tidak bisa disebutkan satu persatu, saya ucapkan terima kasih yang sebesar-besarnya.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ABSTRAK

Ulasan Pengguna pada platform e-commerce merupakan sumber informasi penting yang mencerminkan tingkat kepuasan maupun ketidakpuasan terhadap layanan dan produk. Seiring meningkatnya jumlah ulasan, proses analisis sentimen secara manual menjadi tidak efektif dan memerlukan waktu yang lama. Oleh karena itu diperlukan pendekatan otomatis untuk mengelompokkan sentimen ulasan secara efisien dan akurat. Penelitian ini bertujuan untuk mengklasifikasikan sentimen ulasan pengguna aplikasi Tokopedia menjadi dua kelas, yaitu positif dan negatif, menggunakan metode Naive Bayes Classifier. Teknik pengumpulan data menggunakan *web scraping*, total data yang diperoleh 3.000 ulasan, kemudian dilakukan pelabelan manual dan tahapan text preprocessing meliputi cleaning, case folding, tokenizing, normalisasi, stopword removal, dan stemming. Pembobotan kata dilakukan menggunakan TF-IDF dengan seleksi fitur threshold 0,00001. Kemudian data diuji menggunakan tiga skenario pembagian data latih-data uji, yaitu 90:10, 80:20, 70 :30. Berdasarkan hasil penelitian didapatkan bahwa model Naive Bayes Classifier mampu memberikan performa yang konsisten dengan akurasi diatas 91%. Nilai akurasi yang diperoleh pada skenario 90:10 adalah 93,67%, pada skenario 80:20 sebesar 93,00%, dan pada skenario 70:30 sebesar 92,56%. Skenario terbaik pada (80:20), model mencapai precision 97,01%, Recall 88,44%, F1-Score 92,53%, serta nilai AUC 0,9787. Hal tersebut menunjukkan bahwa metode Naive Bayes Classifier terbukti efektif dalam mengklasifikasikan sentimen ulasan pengguna Tokopedia, dengan performa konsisten dan tingkat akurasi yang tinggi.

Kata kunci: Analisis sentimen, Google Play, Klasifikasi, Naive Bayes Classifier, Tokopedia.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ABSTRACT

User reviews on e-commerce platforms are an important source of information that reflects the level of satisfaction or dissatisfaction with services and products. As the number of reviews increases, manual sentiment analysis becomes ineffective and time-consuming. Therefore, an automated approach is needed to group review sentiments efficiently and accurately. This study aims to classify the sentiments of Tokopedia app user review into two classes, positive and negative, using the Naive Bayes Classifier method. The data collection technique used web scraping, with a total of 3,000 reviews obtained, followed by manual labeling and text preprocessing stages including cleaning, case folding, tokenizing, normalization, stopword removal, and stemming. Word weighting was performed using TF-IDF with a feature selection threshold of 0,00001. The data was then tested using three training-testing data distribution scenarios, namely 90:10, 80:20, 70:30. Based on the results of the study, it was found that the Naive Bayes Classifier model was able to provide consistent performance with an accuracy of above 91%. The accuracy value obtained in the 90:10 scenario was 91,67%, in the 80:20 scenario was 93,00%, and in the 70:30 scenario was 92,56%. In the best scenario (80:20), the model achieved a precision of 97,01%, a recall of 88,44% an F1-Score of 92,53%, and AUC value of 0,9787. This shows that the Naive Bayes Classifier method is effective in classifying the sentiment of Tokopedia user reviews, with consistent performance and a high level of accuracy.

Keyword : Sentiment Analysis, Google Play, Classification, Naive Bayes Classifier, Tokopedia.



Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Assalammu'alaikum wa rohmatullohi wa barokatuh.

Alhamdulillah robbil'alamin, kami bersyukur kepada Allah SWT atas karunia dan petunjuk yang diberikan kepada penulis sehingga laporan tugas akhir ini dapat diselesaikan dengan sukses. "Klasifikasi Sentimen Aplikasi Tokopedia menggunakan pendekatan Naive Bayes Classifier berdasarkan ulasan di Google Play" adalah judul laporan ini. Ingatlah keberkahan dan salam Rasulullah SAW, yang telah membawa umatnya dari dunia kebodohan ke dunia kebijaksanaan.

Penyelesaian Laporan Tugas Akhir ini merupakan prasyarat untuk program Sarjana Strata 1 (S1) Teknik Informatika di Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim, Riau. Banyak orang yang membantu penulis dalam penulisan laporan ini, baik secara materi maupun inspirasi.

Tugas akhir ini disusun sebagai salah satu syarat untuk mendapatkan gelar sarjana pada jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Banyak yang telah membantu saya dalam menyelesaikan rangkaian penelitian tugas akhir ini, baik dalam bentuk materi, doa maupun motivasi kepada saya. Semua itu terlalu banyak bagi saya untuk membalasnya, namun pada kesempatan ini saya mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Hj. Leny Nofianti, MS, SE, M.Si, Ak. Selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Ibu Dr. Yuslenita Muda, S.Si., M.Sc Selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Bapak Muhammad Affandes, S.T., M.T. Selaku Kepala Prodi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Bapak Dr. Alwis Nazir, M.Kom selaku dosen pembimbing akademik yang telah memberikan nasehat serta semangat dan membantu saya dalam proses belajar selama di Jurusan Teknik Informatika.
5. Ibu Yusra, S.T, M.T dan Bapak Muhammad Fikry, S.T, M.Sc. selaku



Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dosen pembimbing TA yang telah memberikan bimbingan berupa arahan dan nasihat kepada penulis sehingga dapat menyelesaikan proposal TA. dan Penguji Ta.

6. Kedua orang tua tercinta, Ayahanda Harianto dan Ibunda Irawati yang selalu memberikan dukungan moril, materil dan tentunya do'a yang tiada hentinya.
7. Adek Satu-satunya yang paling disayangi yaitu Ressya Anggun Apriyani, yang telah memberikan doa, dukungan dan semangat yang sangat luar biasa kepada penulis dalam menyelesaikan proposal TA ini.
8. Seluruh teman-teman di Jurusan Teknik Informatika UIN Suska Riau angkatan 2020 terkhusus untuk kelas I yang bersama-sama sedang berjuang meraih gelar S.T.
9. Seluruh pihak yang belum disebutkan satu-satu penulis mengucapkan terima kasih atas dukungan yang telah diberikan secara material maupun spiritual.

Penulis mengakui bahwa terdapat banyak kesalahan dan kekurangan dalam penelitian ini. Oleh karena itu, kami terbuka terhadap rekomendasi untuk perbaikan dan kritik yang membangun. Terakhir, kami berharap setiap orang yang membaca laporan ini akan merasa bermanfaat.

Wassalamu'alaikum wa rohmatullohi wa barokatuh.

Pekanbaru, 17 November 2025

UIN SUSKA RIAU

Penulis



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR ISI

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL	ii
LEMBAR PERNYATAAN.....	iii
LEMBAR PERSEMBAHAN	iv
ABSTRAK.....	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiv
DAFTAR RUMUS	xv
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
BAB 2 KAJIAN PUSTAKA	4
2.1 Kajian Metode	4
2.1.1 Text Mining	4
2.1.2 Opinion Mining.....	5
2.1.3 Teks Preprocessing.....	5
2.1.4 Naïve Bayes Classifier	6
2.1.5 Confusion Matrix.....	8
2.2 Perbandingan Penelitian ini dengan Penelitian sebelumnya.....	8
BAB 3 METODOLOGI PENELITIAN	10
3.1 Tahapan Penelitian.....	10
3.2 Identifikasi Masalah.....	11
3.3 Pengumpulan Data	11
3.4 Pelabelan.....	12

3.5	Text Preprocessing.....	12
3.5.1	Cleaning	13
3.5.2	Case Folding	13
3.5.3	Tokenizing	13
3.5.4	Normalisasi	14
3.5.5	Negation Handling	14
3.5.6	Stopword Removal	15
3.5.7	Stemming.....	15
3.6	Pembobotan	16
3.7	Analisis Metode	16
3.8	Pengujian	17
3.8.1	Confusion Matrix.....	18
3.8.2	Metrik Evaluasi.....	19
3.8.3	Prosedur Pengujian	19
3.9	Kesimpulan dan Saran	20
BAB IV HASIL DAN PEMBAHASAN.....		21
4.1	Analisa	21
4.1.1	Pengambilan data	21
4.1.2	Labelling	22
4.1.3	Preprocessing Data.....	22
4.1.4	Pembobotan Kata	28
4.1.5	Klasifikasi Naïve Bayes Classifier.....	31
4.2	Implementasi.....	35
4.2.1	Batasan implementasi.....	35
4.2.2	Lingkungan implementasi.....	35
4.2.3	Langkah implementasi	36
4.3	Evaluasi hasil Klasifikasi Pengujian.....	54
4.4	Wordcloud	57
BAB V PENUTUP		58
5.1	Kesimpulan	58
5.2	Saran	58

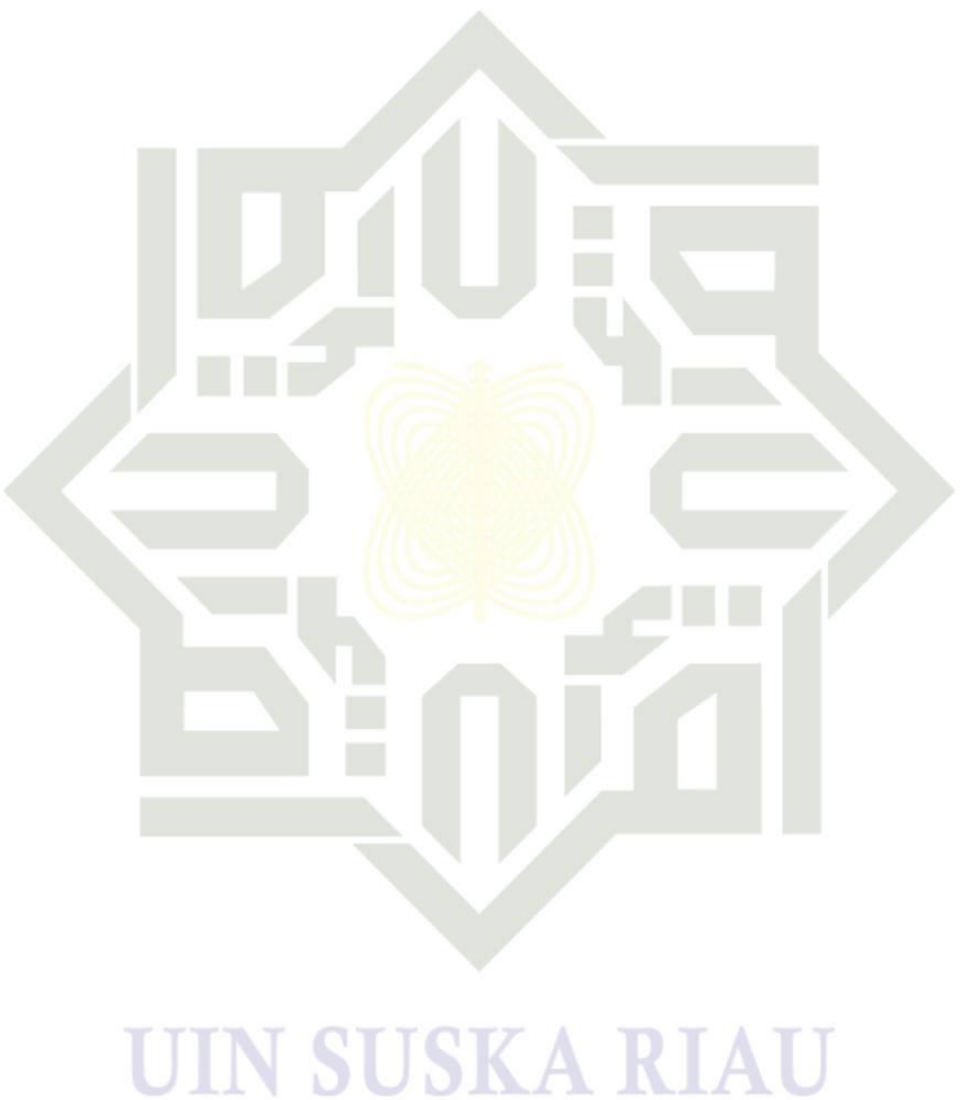
Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA	60
LAMPIRAN A	63
LAMPIRAN B	69

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR GAMBAR

Gambar 3. 1 Tahapan Penelitian	11
Gambar 3. 2 Tahapan text preprocessing	13
Gambar 3. 3 Proses Naïve Bayes Classifier	18
Gambar 4. 1 Contoh Hail Scraping	22
Gambar 4. 2 Hasil Pelabelan	23
Gambar 4. 3 <i>code install library</i>	37
Gambar 4. 4 <i>code import library</i>	37
Gambar 4. 5 <i>Code menjalankan scradi ping</i>	38
Gambar 4. 6 <i>Code untuk menyimpan hasil komentar</i>	38
Gambar 4. 7 Data komentar yang diperoleh dari <i>scraping</i>	39
Gambar 4. 8 Hasil <i>Labelling</i>	39
Gambar 4. 9 <i>Code program untuk tahap cleaning</i>	40
Gambar 4. 10 Output dari proses <i>cleaning</i>	40
Gambar 4. 11 Kode program <i>case folding</i>	41
Gambar 4. 12 Hasil <i>case folding</i>	41
Gambar 4. 13 <i>Code program tahapan tokenizing</i>	41
Gambar 4. 14 Hasil tahapan <i>tokenizing</i>	42
Gambar 4. 15 <i>Code program untuk tahap normalization</i>	42
Gambar 4. 16 Output dari proses <i>normalization</i>	43
Gambar 4. 17 <i>Code program untuk tahapan negation handling</i>	43
Gambar 4. 18 Output dari proses <i>negation handling</i>	44
Gambar 4. 19 <i>Code program untuk tahapan stopword removal</i>	44
Gambar 4. 20 Output dari <i>stopword removal</i>	45
Gambar 4. 21 <i>Code program untuk stemming</i>	45
Gambar 4. 22 hasil tahapan <i>stemming</i>	46
Gambar 4. 23 <i>code TF-IDF</i>	46
Gambar 4. 24 <i>code thresholding</i>	47
Gambar 4. 25 contoh kata <i>selected features</i>	47
Gambar 4. 26 Data <i>frame TF-IDF</i>	47



Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Gambar 4. 27 <i>code training</i>	48
Gambar 4. 28 <i>code prediksi</i>	48
Gambar 4. 29 hasil klasifikasi 90% : 10%.....	48
Gambar 4. 30 hasil klasifikasi 80% : 20%.....	49
Gambar 4. 31 hasil klasifikasi 70% : 30%.....	49
Gambar 4. 32 <i>confusion matrix</i> perbandingan data latih 90% dan data uji 10%	50
Gambar 4. 33 <i>confusion matrix</i> perbandingan data latih 80% dan data uji 20%	50
Gambar 4. 34 <i>confusion matrix</i> perbandingan data latih 70% dan data uji 30%	51
Gambar 4. 35 tanpa Threshold	54
Gambar 4. 36 Perbandingan Threshold 0,001	55
Gambar 4. 37 Perbandingan Threshold 0,00001	55
Gambar 4. 38 Wordcloud Negatif	58
Gambar 4. 39 Wordcloud Positif.....	58



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR TABEL

Tabel 2. 1 Penelitian analisis sentimen yang menggunakan metode NBC.....	4
Tabel 3. 1 Contoh Pembersihan data.....	14
Tabel 3. 2 Contoh Case Folding.....	14
Tabel 3. 3 Contoh Tokenizing.....	15
Tabel 3. 4 Contoh Normalisasi.....	15
Tabel 3. 5 Contoh Stopword Removal.....	16
Tabel 3. 6 Contoh Stemming.....	17
Tabel 3. 7 Confusion Matrix.....	19
Tabel 4. 1 Hasil <i>Cleaning</i>	24
Tabel 4. 2 Hasil <i>Case Folding</i>	25
Tabel 4. 3 Output dari Tokenisasi.....	26
Tabel 4. 4 Output dari <i>Normalization</i>	27
Tabel 4. 5 Output dari <i>Negation Handling</i>	28
Tabel 4. 6 Output dari Penghapusan Stopword.....	29
Tabel 4. 7 Output dari <i>Stemming</i>	29
Tabel 4. 8 Pembobotan pada Kata.....	31
Tabel 4. 9 Probabilitas Kata.....	33
Tabel 4. 10 Sampel pada Data Testing.....	34
Tabel 4. 11 Data Pengujian sesudah preprocessing.....	34
Tabel 4. 12 Sampel data pelatihan dan data pengujian.....	34
Tabel 4. 13 Data Uji.....	35
Tabel 4. 14 Threshold 90:10.....	52
Tabel 4. 15 Threshold 80:20.....	53
Tabel 4. 16 Threshold 70:30.....	53
Tabel 4. 17 Hasil Confusion matrix.....	56
Tabel 4. 18 Nilai Accuracy.....	57



Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR RUMUS

(1) Naive Bayes Classifier.....	6
(2) Probabilitas Bersyarat.....	7
(3) Probabilitas Kelas Positif.....	7
(4) Probabilitas Kelas Negatif.....	7
(5) Probabilitas setiap kata.....	7
(6) Akurasi.....	8
(7) Presisi.....	8
(8) Recall.....	8
(9) F1-Score.....	8
(10) AUC.....	8
(11) TPR.....	8
(12) FPR.....	8

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Ulasan pengguna *e-commerce* merupakan metode yang cepat dan tepat untuk memperoleh informasi yang diberikan berhubungan dengan produk, beberapa penelitian menunjukkan bahwa hampir setengah mereka yang menggunakan internet bergantung pada saran dari *word of mouth* sebelum membeli barang apapun. Ulasan dan komentar pengguna lain memberikan informasi terbaru dan perspektif yang berguna bagi calon pembeli (Fanissa et al., 2018). Pengguna yang tidak puas dengan layanan atau produk perusahaan biasanya menulis komentar keluhan di media sosial. Di sisi lain, beberapa pengguna merasa puas dengan produk dan mengungkapkan kesenangan mereka melalui komentar di media sosial. Komentar pengguna di media sosial meskipun sedikit, dapat memengaruhi calon pelanggan (Joesyiana, 2018). Namun memantau dan mengatur komentar dari masyarakat juga tidaklah mudah.

Studi tentang pandangan publik, sentimen, evaluasi, penilaian, sikap, serta emosi terhadap suatu hal, misalnya produk, layanan, perusahaan, individu, masalah, peristiwa, topik, dan atributnya, dikenal sebagai analisis sentimen, atau penambangan opini. (Juniarsih et al., 2020). Namun, bersamaan dengan jumlah ulasan yang semakin banyak, proses analisis secara manual menjadi kurang efektif dan membutuhkan waktu yang cukup lama.

Untuk mengatasi permasalahan tersebut, digunakan pendekatan otomatis dalam pengolahan dan pengelompokan ulasan. Pendekatan yang digunakan adalah klasifikasi sentimen, yaitu pengklasifikasian ulasan ke dalam sentimen positif dan negatif melalui pendekatan *machine learning* (Larasati, 2025). Dengan demikian, fokus penelitian ini terletak pada penerapan klasifikasi sentimen secara otomatis bukan hanya sekedar membahas konsep sentimen.

Penelitian ini menggunakan metode Naïve Bayes Classifier yang umum diterapkan dalam analisis sentimen. Metode klasifikasi yang memanfaatkan



Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

perhitungan peluang untuk memprediksi kemungkinan sebuah peristiwa dengan melihat jumlah kemunculan tiap kelas pada data training disebut Naive Bayes Classifier (Rahel Lina Simanjuntak et al., 2023). Naïve Bayes Classifier sendiri memiliki keunggulan berupa penerapan dan optimalisasi dalam penggunaan tenaga pemrosesan dibanding melalui metode lainnya. Naïve bayes juga mampu menunjukkan performa yang sangat baik dalam tugas klasifikasi teks, kesederhanaan, dan efisiensi khususnya di berbagai aplikasi praktis seperti spam filtering atau klasifikasi artikel berita (Aditya Nugroho & Cholissodin, 2021).

Pendekatan Naive Bayes Classifier telah banyak digunakan dalam studi-studi sebelumnya untuk analisis sentimen, khususnya pada aplikasi Tokopedia. Studi Salsa dkk. menggunakan pendekatan Naive Bayes Classifier untuk meneliti sentimen pelanggan Tokopedia. Penelitian tersebut menggunakan 714 data ulasan dan hanya menerapkan satu skenario pembagian data *training* dan data *testing* yaitu 80:20 dalam pengujian dan menghasilkan akurasi sebesar 95,10% serta nilai AUC sebesar 0,999. Meskipun hasil tersebut menunjukkan performa yang baik, penelitian tersebut belum menerapkan tahapan preprocessing teks secara lengkap serta belum menggunakan pembobotan kata TF-IDF, selain itu penelitian tersebut hanya menggunakan empat tahapan *preprocessing*, yaitu *cleaning*, *filtering*, *tokenizing* dan *stemming*. (Melina Salsabila et al., 2022).

Berdasarkan penelitian tersebut, ada peluang untuk meningkatkan kualitas proses klasifikasi sentimen melalui penggunaan dataset yang lebih besar, penerapan preprocessing teks yang lebih lengkap serta penggunaan pembobotan kata. Oleh karena itu, penelitian ini menggunakan 3.000 ulasan Tokopedia yang diperoleh dari Google PlayStore, menerapkan tahapan preprocessing yang lebih lengkap mencakup *cleaning*, *case folding*, *tokenisasi*, *normalisasi*, *penanganan negasi*, *penghapusan stopword* dan *stemming*. Selain itu, mengimplementasikan pembobotan TF-IDF disertai seleksi fitur yang ditentukan oleh nilai *threshold*, serta melakukan evaluasi performa model melalui beberapa skenario pembagian data *training* dan data *testing*, yaitu



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

90:10. 80:20, 70:30. Dengan pendekatan tersebut. Melalui penelitian ini diharapkan diperoleh model klasifikasi sentimen yang lebih konsisten dan akurat dalam menggambarkan persepsi pengguna terhadap aplikasi Tokopedia.

1.2 Rumusan Masalah

Berdasarkan pendahuluan yang telah dipaparkan, rumusan masalah dalam tugas akhir ini adalah:

1. Bagaimana penerapan metode Naïve Bayes Classifier dalam melakukan klasifikasi sentimen pada ulasan aplikasi Tokopedia di Google Play?
2. Bagaimana performa metode Naïve Bayes Classifier berdasarkan beberapa skenario pembagian data latih dan data uji dalam klasifikasi sentimen ulasan aplikasi Tokopedia di Google Play?

1.3 Batasan Masalah

Beberapa batasan masalah pada penelitian ini yaitu:

1. Data ulasan yang dipakai berjumlah 3.000 data dan hanya akan diklasifikasi sebagai data ulasan pengguna Tokopedia di Google PlayStore.
2. Review dalam aplikasi yang akan diklasifikasikan harus bahasa Indonesia.
3. Model membagi klasifikasi menjadi dua kelas, yaitu positif dan negatif.

1.4 Tujuan Penelitian

Tujuan penelitian dalam Tugas Akhir ini meliputi:

1. Menerapkan metode Naive Bayes Classifier untuk melakukan klasifikasi sentimen pada ulasan aplikasi Tokopedia di Google Play.
2. Menguji performa metode Naive Bayes Classifier menggunakan beberapa skenario pembagian data latih dan data uji.

1.5 Manfaat Penelitian

Dengan mengklasifikasikan sentimen pengguna ke dalam kategori positif dan negatif, penelitian ini bertujuan untuk menjelaskan bagaimana perasaan pengguna terhadap aplikasi Tokopedia berdasarkan ulasan di Google Play Store. Selain itu, temuan penelitian ini dapat menjadi referensi untuk penelitian selanjutnya yang menggunakan Naïve Bayes Classifier untuk kategorisasi sentimen, khususnya pada aplikasi e-commerce Indonesia.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 2 KAJIAN PUSTAKA

2.1 Kajian Metode

Studi terhadap metode penelitian membantu peneliti menentukan pendekatan yang paling relevan serta sesuai dengan tujuan studi. Selain itu, kajian ini memperkaya pemahaman mengenai metodologi dengan menghimpun berbagai penjelasan yang berhubungan dengan topik penelitian.

2.1.1 Text Mining

Pada proses pengolahan data, Text Mining tergolong dalam pengolahan data yang berfokus pada pengenalan struktur dalam teks, teks mining memanfaatkan data tidak terstruktur yang menjadi pembeda utama dibandingkan dengan teks mining yang menggunakan data terstruktur atau basis data sebagai masukan. Tujuannya untuk memperoleh informasi berharga dari dokumen (Wibowo et al., 2024)

Sama halnya Data mining, Text mining juga digunakan untuk mengekstrak informasi dan wawasan berharga dari kumpulan dataset dalam jumlah besar. Meskipun keduanya dapat menggunakan data yang berbentuk database, mereka memiliki tipe data yang berbeda. Data Mining biasanya bekerja dengan data terstruktur, namun data yang biasanya digunakan untuk text mining adalah data yang tidak terstruktur maupun semi terstruktur. yang menimbulkan masalah tambahan. seperti kerumitan dan ketidaklengkapan struktur teks, makna yang ambigu, serta variasi bahasa dan kesalahan dalam terjemahan (Setiawan et al., 2025).

Umumnya teks yang diproses dalam teks mining terdapat sejumlah karakteristik, misalnya tergolong berdimensi tinggi, yang berarti mengandung gangguan pada data dan juga terdapat data atau keterangan yang tidak relevan dan tidak dibutuhkan, yang berpotensi menghambat proses analisis (Sari1 et al., 2020). Untuk menganalisis sentimen emosional dan menentukan apakah pernyataan tersebut positif atau negatif, penambahan teks dapat menawarkan informasi yang bermanfaat dari sumber opini publik.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.1.2 Opinion Mining

Opinion Mining adalah bagian dari teks mining yang berfokus pada analisis opini atau sudut pandang yang ditemukan dalam dokumen teks. Penambangan opini adalah cabang studi yang meneliti pikiran, perasaan, penilaian, sikap, dan emosi seseorang terhadap entitas dan karakteristiknya sebagaimana tercermin pada teks tertulis (Prabowo & Wiguna, 2021).

2.1.3 Teks Preprocessing

Pra-pemrosesan teks merupakan langkah pertama dalam pengolahan teks yang mengubah teks menjadi format yang dapat diproses lebih lanjut. Teks ini perlu dipecah menjadi beberapa tingkatan, Dokumen dapat diuraikan menjadi: Bab, Bagian, Paragraf, Kalimat, blok kata/karakter, selain itu langkah ini menghilangkan dan mengubah tampilan angka, huruf kapital, dan tanda lainnya (Anggi Pranata, 2022).

Berikut langkah-langkah *preprocessing* yang didasarkan pada prosesnya (Sriwenda Putri et al., 2023):

1. Cleaning

Proses pembersihan data melibatkan penghapusan data yang tidak konsisten, berlebihan, tidak valid, dan tidak relevan. Dengan mengurangi kuantitas dan kompleksitas data, prosedur ini juga meningkatkan kinerja penambangan data (Rizki Rinaldi et al., 2024).

2. Case Folding

Berupa langkah untuk mengubah semua huruf dalam teks menjadi huruf kecil atau huruf besar, agar tidak ada perbedaan berdasarkan kapitalisasi dalam penguraian teks. Misalnya, “Halo” dan “halo” akan dianggap sama setelah proses case folding (Hasanah & Sari, 2024).

3. Stopword Removal

Yakni mengurangi jumlah kata dengan menghilangkan stopwords. Stopword merujuk pada kata-kata yang tidak unik atau spesifik pada suatu dokumen, contohnya kata-kata penghubung atau kata-kata kepemilikan.

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4. Normalization

Suatu proses penormalisasi kosakata yang awalnya tidak baku menjadi bentuk kata yang baku. Contohnya, menggantikan “udah” menjadi “sudah” atau “gimana” menjadi “bagaimana”.

5. Negation Handling

Penanganan negasi adalah proses mengenali dan memproses kata-kata yang membalikkan makna sentimen dari kata atau frasa di dekatnya,

6. Tokenizing

Berupa proses membagi dokumen atau teks menjadi token, yang biasanya b kata-kata atau simbol-simbol individu yang dalam keadaan tertentu memiliki arti atau makna tertentu (Aida et al., 2024).

7. Stemming

Proses ini bertujuan melakukan perubahan pada kata diubah ke bentuk dasar dan menghapus imbuhan atau afikasi pada kata dalam dokumen atau mengubah kata kerja menjadi kata benda. Sebagai contoh, kata “dihilangkan” diubah menjadi “hilang” setelah menghapus imbuhan “di-” dan “kan” (Putra et al., 2020).

2.1.4 Naïve Bayes Classifier

Naïve bayes Classifier digunakan karena termasuk ke dalam metode klasifikasi sederhana dan efektif. Kemampuan klasifikasi teori Naive Bayes serupa dengan kemampuan neural network dan decision tree, dan dapat diterapkan pada data yang sangat besar atau tidak lengkap. Saat digunakan pada basis data berukuran sangat besar, Algoritma Naïve Bayes sangat cepat dan akurat (Fitrana et al., 2024). Model Naïve Bayes dinyatakan dengan persamaan seperti berikut:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

Keterangan:

$P(A|B)$ = Peluang terjadinya hipotesis A ketika diketahui bahwa bukti B muncul.

$P(B|A)$ = Peluang munculnya bukti B dapat memengaruhi

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

kebenaran hipotesis A

$$\begin{aligned}
 P(A) &= \text{Peluang awal terjadinya hipotesis A tanpa mempertimbangkan bukti apapun} \\
 P(B) &= \text{Peluang awal terjadinya bukti B tanpa mempertimbangkan bukti lainnya.}
 \end{aligned}$$

Baik untuk klasifikasi biner maupun multikelas, metode Naive Bayes berfungsi dengan baik. Ini dikenal sebagai klasifikasi Naive Bayes dan menggunakan teknik klasifikasi objek observasi untuk memprediksi label kelas berdasarkan data yang tidak diketahui menggunakan probabilitas bersyarat. Likelihood adalah ukuran kemungkinan terjadinya suatu peristiwa berdasarkan peristiwa atau kejadian lain yang diketahui, dengan menggunakan asumsi, hipotesis, informasi, dan aturan dasar. Berikut tahapan algoritma Naive Bayes Classifier:

1. Menghitung probabilitas bersyarat/likelihood:

$$P(x|C) = P(x_1, x_2, \dots, x_n|C) \quad (2)$$

Keterangan:

C = Kelas

x = vektor dari nilai atribut n

$P(x_i | C)$ = proporsi dokumen dari kelas C yang mengandung nilai atribut x_i

- a. Menghitung probabilitas prior untuk setiap kelas

$$p(\text{positif}) = \frac{fd(\text{positif})}{|D|} = \frac{2}{4} = 0,5 \quad (3)$$

$$p(\text{negatif}) = \frac{fd(\text{negatif})}{|D|} = \frac{2}{4} = 0,5 \quad (4)$$

- b. Menghitung nilai peluang setiap kata:

$$p(x_k|V_j) = \frac{n_k + 1}{|n + \text{kosakata}|} \quad (5)$$



Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.1.5 Confusion Matrix

Tahapan selanjutnya adalah melakukan evaluasi. Metode evaluasi yang digunakan untuk mengevaluasi metode klasifikasi disebut dengan *Confusion Matrix*. Empat istilah menggambarkan hasil klasifikasi dalam pengukuran efisiensi yaitu True Positif (TP), True Negative (TN), False Positive (FP), serta False Negative (FN) (Normawati & Prayogi, 2021). Rumus berikut dapat digunakan untuk menghitung akurasi, presisi dan recall setelah mengumpulkan nilai dari keempat elemen tersebut:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

$$\text{F1 Score} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (9)$$

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR}) d(\text{FPR}) \quad (10)$$

$$\text{TPR} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{FPR} = \frac{FP}{FP + TN} \quad (12)$$

2.2 Perbandingan penelitian ini dengan penelitian sebelumnya

Berdasarkan hasil perbandingan dengan penelitian sebelumnya yang dilakukan oleh Sifa Melina Salsabila et al., penelitian ini memiliki beberapa kelebihan yang menunjukkan adanya pengembangan dari penelitian terdahulu. Kelebihan utama penelitian ini terletak pada penggunaan dataset yang lebih besar, yaitu sekitar 3.000 ulasan pengguna aplikasi Tokopedia, dibandingkan penelitian sebelumnya yang hanya menggunakan 714 ulasan. Penggunaan jumlah data yang lebih besar membuat data yang dianalisis menjadi lebih representatif terhadap opini pengguna, sehingga hasil klasifikasi yang diperoleh diharapkan lebih mencerminkan kondisi sebenarnya.

Selain jumlah data, variasi juga dapat dilihat pada teknik pra-pemrosesan



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

yang digunakan. Hanya teknik pra-pemrosesan paling mendasar—pembersihan, penyaringan, tokenisasi, dan stemming—yang digunakan dalam penelitian sebelumnya. Teknik pra-pemrosesan yang lebih menyeluruh, seperti pembersihan, case folding, tokenisasi, normalisasi kata non-standar, penanganan negasi, penghapusan stopword, dan stemming, digunakan dalam penelitian ini. Dengan mengurangi noise dan meningkatkan struktur kata, metode pra-pemrosesan ini bertujuan untuk meningkatkan kualitas input teks, sehingga memungkinkan prosedur klasifikasi yang lebih efektif.

Penelitian sebelumnya menggunakan pendekatan probabilistik dalam metode Naive Bayes tanpa menerapkan pembobotan khusus untuk fitur. Sebaliknya, penelitian ini menyaring kata-kata dengan bobot rendah dan sedikit pengaruh pada proses klasifikasi menggunakan pemilihan fitur berbasis ambang batas bersamaan dengan pembobotan TF-IDF. Kuantitas karakteristik yang digunakan lebih efisien dan relevan berkat pemilihan fitur ini. Kinerja model hanya dinilai menggunakan satu kondisi pengujian karena penelitian sebelumnya hanya menggunakan satu skenario pembagian data selama tahap penilaian: 80% data pelatihan dan 20% data pengujian. Berbeda dengan penelitian sebelumnya, penelitian ini menunjukkan stabilitas kinerja model di berbagai rasio data pelatihan dan pengujian menggunakan tiga skenario pembagian data: 90:10, 80:20, dan 70:30. Hal ini memberikan pemahaman yang lebih lengkap tentang stabilitas metode tersebut.

Namun, terdapat sejumlah kekurangan dalam penelitian ini. Perasaan netral dan emosi yang lebih bernuansa lainnya tidak dapat diakomodasi oleh sistem klasifikasi sentimen, yang masih terbatas pada dua kelas: positif dan negatif. Selain itu, penelitian ini hanya menggunakan Naive Bayes Classifier dan tidak membandingkannya dengan teknik klasifikasi lain seperti Random Forest atau Support Vector Machines. Selanjutnya, prosedur pelabelan data dilakukan secara manual, yang dapat menimbulkan subjektivitas dari peneliti. Oleh karena itu, meskipun metodologi dan sistem penilaian dalam penelitian ini lengkap, masih ada ruang untuk perbaikan dalam penelitian selanjutnya, terutama dalam hal menambahkan kelas sentimen dan membandingkannya dengan teknik kategorisasi alternatif.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 3

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Serangkaian prosedur yang diambil sepanjang proses pelaksanaan penelitian guna memastikan kelancaran, struktur, dan pencapaian tujuan yang diinginkan disebut dengan metodologi penelitian. Metode penelitian ini memiliki tahapan seperti berikut:



Gambar 3. 1 Tahapan Penelitian

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.2 Identifikasi Masalah

Bagian berikut adalah proses awal dalam kerangka metodologi digunakan untuk mengidentifikasi Pendahuluan, masalah yang diteliti, dan menetapkan tujuan penelitian. Dalam langkah ini, batasan-batasan penelitian juga ditentukan. Terkait hal tersebut maka langkah-langkah yang diterapkan dalam pada tahap identifikasi masalah:

a. Latar Belakang

Berupa dasar yang mendorong perlunya melakukan penelitian, dalam hal ini fenomena Aplikasi Tokopedia di Google PlayStore serta kebiasaan pengguna dalam memberikan tinjauan mengenai Aplikasi Tokopedia di Google PlayStore menjadi dasar penelitian ini.

b. Tujuan Penelitian

Mengklasifikasikan Analisis Sentimen dengan sangat akurat ke dalam kategori positif dan negatif.

c. Rumusan Masalah

Berupa aspek-aspek yang perlu diteliti, seperti pemilihan metode dan pengumpulan data komentar dari Aplikasi Tokopedia di Google PlayStore.

d. Batasan Masalah

Berupa tahap yang penting dalam identifikasi diperlukan agar penelitian dapat lebih terfokus dan mencapai tujuan yang ditetapkan.

3.3 Pengumpulan Data

Dalam langkah mengumpulkan data ini, informasi bersumber dari Google PlayStore Tokopedia yang dapat diakses menggunakan perangkat android atau melalui browser dikomputer. Data yang dikumpulkan dari ulasan pengguna, yaitu:

- a. Teks Ulasan: Komentar yang ditulis oleh pengguna mengenai pengalaman mereka menggunakan aplikasi dan ulasan menggunakan Bahasa Indonesia.
- b. Rating: skor yang diberikan oleh pengguna
- c. Tanggal Ulasan: Tanggal ketika ulasan ditulis.
- d. Nama pengguna: nama atau inisial pengguna (jika tersedia)

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pada tahap Scraping, proses dilakukan dengan menjalankan kode Python melalui Google Colab. Dari hasil scraping terkumpul sebanyak 3.000 ulasan pengguna, mencakup kelas positif dan negatif dari periode waktu yang berbeda (Mei 2024 – Agustus 2025) setelah pengumpulan data selesai, dilakukan pelabelan manual.

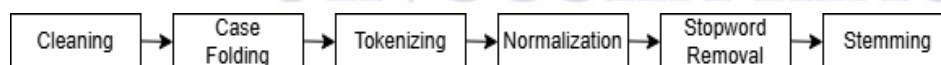
3.4 Pelabelan

Setelah data berhasil dikumpulkan dengan web scraping tahap berikutnya pelabelan (*labelling*). Proses pelabelan terhadap data yang sudah dikumpulkan sebelumnya disebut dengan *labelling*. Masing-masing data diberikan label positif dan negatif sesuai dengan arti dari satu pernyataan pada dataset tersebut. Pelabelan dilaksanakan pada bulan Agustus 2025 tepat setelah proses *Scraping* selesai pada awal bulan yang sama, sehingga memakan waktu hampir sebulan lebih untuk menyelesaikan seluruh dataset. Jumlah data yang dilabel sebanyak 3000 ulasan pengguna, yaitu 1500 data positif dan 1500 data negatif.

Data dilabeli secara manual oleh peneliti dan dilakukan validasi oleh dosen Bahasa Indonesia, Elvina, M.Pd. pelabelan manual dipilih karena mampu menangkap makna tersirat dalam ulasan berbahasa Indonesia dibandingkan metode otomatis. Data hasil pelabelan disimpan dalam format csv yang terdiri dari kolom teks ulasan dan label.

3.5 Text Preprocessing

Berbagai strategi dan metode yang diterapkan agar manipulasi juga memproses teks melalui metode komputasi disebut Teks Preprocessing. Tahapan ini bertujuan untuk memperbaiki format data agar konsisten dan tepat sehingga bisa digunakan untuk analisis.



Gambar 3. 2 Tahapan text preprocessing

Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.5.1 Pembersihan Data

Dalam fase ini, dilakukan penghilangan karakter khusus beserta karakter yang tidak diperlukan dalam teks seperti tanda baca, URL, tag HTML dan elemen lainnya.

Tabel 3. 1 Contoh dari Pembersihan data

Sebelum Cleaning	Sesudah Cleaning
Seneng banget blanja di tokopedia, apalagi kalo banyak promo diskon nya.. belanja bisa buat dijual lagi.. jdi dapet cuan..	Seneng banget blanja di tokopedia apalagi kalo banyak promo diskon nya belanja bisa buat dijual lagi jdi dapet cuan

3.5.2 Case Folding

Pengubahan huruf menjadi huruf kecil adalah proses mengubah setiap huruf dalam teks menjadi huruf kecil untuk menghilangkan potensi perbedaan dalam interpretasi teks akibat penggunaan huruf kapital.

Tabel 3. 2 Contoh Case Folding

Sebelum Case Folding	Sesudah Case Folding
Seneng banget blanja di tokopedia, apalagi kalo banyak promo diskon nya.. belanja bisa buat dijual.. lagi jdi dapet cuan..	seneng banget blanja di tokopedia apalagi kalo banyak promo diskon nya belanja bisa buat dijual lagi jdi dapet cuan

3.5.3 Tokenizing

Tokenisasi merupakan suatu proses di mana teks ulasan Tokopedia dibagi menjadi bagian-bagian kecil yang disebut token. Token ini umumnya berupa kata dan berfungsi sebagai unit dasar yang mewakili informasi dalam teks.

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 3. 3 Contoh Tokenizing

Sebelum Tokenizing	Sesudah Tokenizing
Seneng banget blanja di tokopedia, apalagi kalo banyak promo diskon nya.. belanja bisa buat dijual lagi.. jdi dapet cuan..	["seneng", "banget", "blanja", "di", "tokopedia", "apalagi", "kalo", "banyak", "promo", "diskon", "nya", "belanja", "bisa", "buat", "dijual", "lagi", "jdi", "dapet", "cuan"]

3.5.4 Normalisasi

Dalam langkah ini, dilakukan pernormalisasi kosakata yang awalnya belum baku diubah menjadi bentuk kata yang baku. Proses ini menggunakan kamus normalisasi yang dibuat manual dalam format csv. Normalisasi dilakukan dengan mencocokkan setiap token dengan daftar kata tidak baku pada kamus normalisasi, kemudian menggantinya untuk menghasilkan data yang konsisten dan siap dianalisis.

Tabel 3. 4 Contoh Normalisasi

Sebelum Normalisasi	Sesudah Normalisasi
["seneng", "banget", "blanja", "di", "tokopedia", "kalo", "banyak", "promo", "nya", "belanja", "bisa", "buat", "dijual", "lagi", "jadi", "dapet", "cuan"]	["senang", "sekali", "belanja", "di", "tokopedia", "kalau", "banyak", "promo", "diskon", "nya", "belanja", "bisa", "buat", "dijual", "lagi", "jadi", "dapat", "cuan"]

3.5.5 Negation Handling

Strategi yang diterapkan untuk mengelola kata berunsur negatif dalam dokumen ulasan disebut dengan negation handling. Seperti "tidak", "bukan", "ngga", "jangan" dan lainnya.

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 3. 5 Contoh Negasi

Sebelum Negation Handling	Sesudah Negation Handling
["senang", "sekali", "belanja", "di", "tokopedia", "apalagi", "kalau", "banyak", "promo", "diskon", "nya", "belanja", "bisa", "buat", "dijual", "lagi", "jadi", "dapat", "cuan"]	["senang", "sekali", "belanja", "di", "tokopedia", "apalagi", "kalau", "banyak", "promo", "diskon", "nya", "belanja", "bisa", "buat", "dijual", "lagi", "jadi", "dapat", "cuan"]

3.5.6 Stopword Removal

Ulasan Tokopedia yang sudah dibersihkan akan dihilangkan karakter, tanda baca, dan kata-kata umum yang tidak relevan atau informatif. Kamus stopwords yang digunakan adalah file csv stopwordsbahasa, kemudian token hasil normalisasi dibandingkan dengan daftar tersebut. Jika ditemukan dalam daftar stopwords, kata tersebut dihapus.

Tabel 3. 6 Contoh Stopword Removal

Sebelum Stopword Removal	Sesudah Stopword Removal
["senang", "sekali", "belanja", "di", "tokopedia", "kalau", "banyak", "promo", "nya", "belanja", "bisa", "buat", "dijual", "lagi", "jadi", "dapat", "cuan"]	["senang", "belanja", "promo", "diskon", "belanja", "dijual", "cuan"]

3.5.7 Stemming

Proses mengubah kata menjadi bentuk dasarnya dengan menghapus imbuhan atau afiksasi pada kata dalam dokumen, atau mengubah kata kerja menjadi kata benda disebut dengan stemming. Pada proses stemming menggunakan library Sastrawi.

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 3. 7 Contoh Stemming

Sebelum Stemming	Sesudah Stemming
Senang banget belanja di tokopedia, apalagi kalau banyak promo diskon nya.. belanja bisa buat dijual lagi.. jadi dapet cuan..	["senang", "belanja", "promo", "diskon", "belanja", "jual", "cuan"]

3.6 Pembobotan

Fitur TF-IDF merepresentasikan setiap dokumen melalui vektor yang terdiri dari elemen-elemen hasil ekstraksi dokumen. Sebuah vektor ini mencakup bobot untuk setiap istilah yang dihitung menggunakan metode dasar TF-IDF. Teknik pembobotan yang frekuensi istilah (TF) dengan menggabungkan frekuensi dokumen invers (IDF) disebut TF-IDF. Setiap kata atau term kemudian diberi nilai bobot dengan TF-IDF, pada penelitian ini, Pembobotan kata dilakukan melalui metode TF-IDF (Term Frequency–Inverse Document Frequency). Setelah nilai TF-IDF diperoleh, dilakukan seleksi fitur menggunakan metode thresholding. Dalam penelitian ini digunakan dua nilai threshold sebagai skenario pengujian, yaitu 0,00001 dan 0,001. proses pembobotan diterapkan pada seluruh dataset yang berjumlah 3000 ulasan bukan hanya contoh ulasan tertentu.

Pentingnya pembobotan TF-IDF terletak pada frekuensi kemunculan kata semakin tinggi nilainya pada suatu dokumen, semakin signifikan kontribusi kata tersebut. Namun, Jika kata tersebut muncul dengan frekuensi lebih tinggi diberbagai dokumen, maka nilai kontribusinya akan lebih kecil (Yutika et al., 2021). Metode (TF-IDF) merupakan suatu cara untuk mengukur sejauh mana sebuah istilah mewakili konten pada teks dengan memberi bobot pada setiap kata di dalamnya (Farah Zhafira et al., 2021).

3.7 Analisis Metode

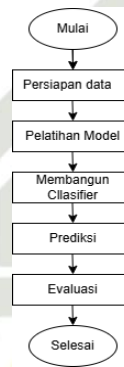
Pendekatan pengklasifikasi Naïve Bayes digunakan dalam fase klasifikasi ini. Model klasifikasi yang sederhana dan efisien adalah pengklasifikasi Naïve Bayes. Studi ini meneliti akurasi dataset yang terdiri dari 3.000 evaluasi pengguna. Data pelatihan dan data pengujian adalah dua kelompok



Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

yang digunakan untuk memisahkan data. Distribusi data eksperimental kemudian diuji dalam tiga cara berbeda: 90% data pelatihan dan 10% data pengujian digunakan dalam eksperimen pertama, 80% data pelatihan dan 20% data pengujian digunakan dalam eksperimen kedua, dan 70% data pelatihan dan 30% data pengujian digunakan dalam eksperimen ketiga.



Gambar 3. 3 Proses Naïve Bayes Classifier

a. Data Training

Pada proses data latih ini, perolehan nilai *prior* dan peluang bersyarat untuk setiap kelas pada data latih digunakan sebagai panduan untuk membuat model klasifikasi.

1. Menentukan nilai peluang bagi setiap kategori (*prior*)
2. Probabilitas untuk setiap term dapat dihitung dengan menggunakan persamaan 9

b. Tahap Uji

Pada langkah ini, data yang diuji dimasukkan kedalam model yang dibangun pada fase pelatihan untuk melakukan pengujian.

3.8 Pengujian

Proses pengujian dilakukan untuk mengevaluasi performa model *Naive Bayes Classifier* dalam mengklasifikasikan ulasan pengguna Tokopedia ke dalam dua kelas, yaitu positif dan negatif. Pengujian ini dilaksanakan setelah proses pelatihan (*training*) selesai serta seluruh data melalui rangkaian preprocessing, pembobotan *TF-IDF*, dan seleksi fitur menggunakan metode



Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Threshold.

Dalam penelitian ini, seleksi fitur berbasis *Threshold TF-IDF* diterapkan sebelum proses pelatihan model dilakukan. Penerapan *Threshold* bertujuan menyaring kata-kata yang memiliki bobot *TF-IDF* sangat rendah karena dianggap tidak memberi kontribusi berarti terhadap proses klasifikasi. Sehingga kata-kata dengan bobot dibawah nilai tersebut dihapus dari kumpulan fitur. Dengan langkah ini, jumlah fitur berkurang noise dapat diminimalkan dan performa model menjadi lebih optimal.

3.3.1 Confusion Matrix

Pada tahap pengujian ini, Banyaknya data atau ulasan yang diklasifikasikan dengan benar atau salah diukur langsung dalam matriks ketidakpastian. Tujuan tabel *Confusion Matrix* yaitu untuk mengevaluasi performa model. Bentuk *Confusion matrix* ditunjukkan seperti berikut:

Tabel 3. 8 Confusion Matrix

		Prediksi Kelas	
		Positif	Negatif
Kelas Sebenarnya	Positif	TP	FN
	Negatif	FP	TN

TP (True Positif) = jumlah dokumen kelas positif yang benar dianggap positif.

TN (True Negatif) = jumlah dokumen kelas negatif yang benar dianggap negatif.

FP (False Positif) = jumlah dokumen kelas negatif yang salah dianggap positif.

FN (False Negatif) = jumlah dokumen kelas positif yang salah dianggap negatif.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.3.2 Metrik Evaluasi

Beberapa metrik digunakan untuk mengevaluasi kualitas model, yaitu:

1. *Accuracy*

Kemungkinan untuk menentukan kelas dengan akurat, mencakup semua nilai yang menjadi bagian dari kelas ini disebut dengan *Accuracy*.

2. *Precision*

Precision merupakan perbandingan data positif yang diklasifikasikan secara tepat terhadap keseluruhan hasil klasifikasi.

3. *Recall*

Recall digunakan untuk mengukur seberapa banyak data positif yang berhasil diklasifikasikan dengan benar dan digunakan untuk mengevaluasi data yang dapat teridentifikasi secara akurat.

4. *F1 Score*

F1 Score merupakan perhitungan rata-rata dari *precision* dan *recall* yang bertujuan menyeimbangkan ketika data yang digunakan tidak seimbang. Nilai *F-1 score* berkisar antara 0-1, ketika nilainya mendekati 1 menunjukkan model klasifikasi yang dihasilkan sudah optimal.

5. *AUC (Area Under Curve)*

AUC berfungsi untuk menilai sejauh mana model mampu membedakan tiap kelas positif dan negatif, Nilai *AUC* menunjukkan seberapa baik model dapat memisahkan dua kelas, nilai yang mendekati 1 menunjukkan performa yang baik. Nilai *AUC* mendekati 1 menunjukkan performa klasifikasi yang sangat baik.

3.3.3 Prosedur Pengujian

Pengujian dilakukan menggunakan tiga skenario pembagian data:

- 70% data latih : 30% data uji
- 80% data latih : 20% data uji
- 90% data latih : 10% data uji

Setiap skenario pengujian menghasilkan nilai confusion matrix, accuracy, precision, recall, *F1-score*, serta *AUC*. Setelah itu, seluruh hasil dari masing-masing skenario dibandingkan untuk menentukan pembagian data yang



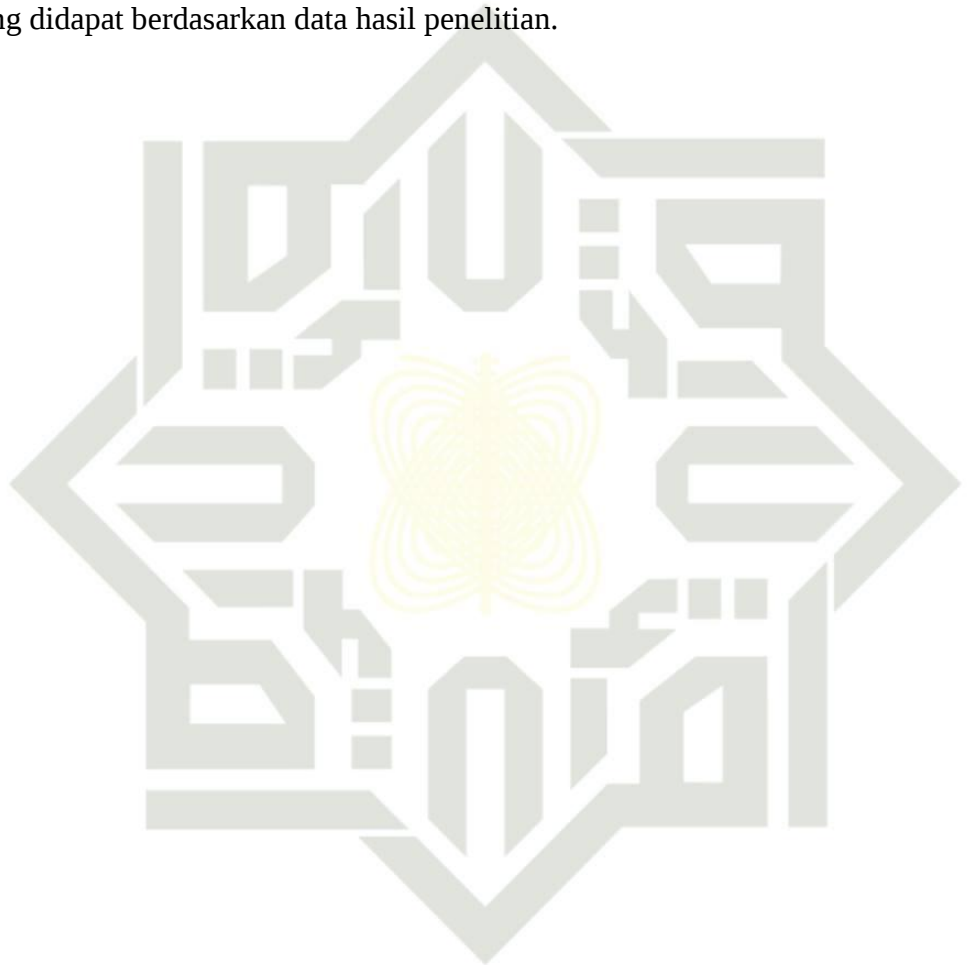
Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

memberikan performa terbaik.

3.9 Kesimpulan dan Saran

Dalam rangka menjamin bahwa penelitian ini sudah dilaksanakan memenuhi kriteria yang telah ditetapkan. Di bagian kesimpulan, ditampilkan hasil evaluasi yang sudah dikerjakan. Rumusan akhir ini berisis poin-poin utama yang didapat berdasarkan data hasil penelitian.



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB V

PENUTUP

5.1 Kesimpulan

1. Sentimen ulasan pengguna aplikasi Tokopedia berhasil diklasifikasikan ke dalam dua kelas sentimen positif dan negatif dengan menggunakan algoritma Naïve Bayes Classifier. Langkah-langkah pra-pemrosesan teks lengkap, termasuk pembersihan, pengubahan huruf besar/kecil, tokenisasi, normalisasi, penanganan negasi, penghapusan stopword, stemming, dan pembobotan kata menggunakan pendekatan TF-IDF dengan pemilihan fitur berbasis ambang batas, membantu proses klasifikasi. Temuan ini menunjukkan bahwa analisis sentimen ulasan berbahasa Indonesia dapat berhasil dilakukan menggunakan pendekatan Naïve Bayes Classifier.
2. Dalam berbagai kondisi pembagian data pelatihan dan pengujian, algoritma Naïve Bayes Classifier secara konsisten dan berhasil menunjukkan kinerja yang baik. Nilai akurasi untuk ketiga skenario pengujian 90:10, 80:20, dan 70:30 semuanya lebih tinggi dari 91%. Pembagian data 80% untuk pelatihan dan 20% untuk pengujian menghasilkan hasil tertinggi, dengan akurasi 93,00%, presisi 97,01%, recall 88,44%, F1-score 92,53%, dan AUC 0,9787. Hal ini menunjukkan bahwa pembagian data tersebut menawarkan keseimbangan terbaik antara pengujian dan pelatihan model.
3. Performa model klasifikasi meningkat secara positif dengan penerapan pembobotan TF-IDF dengan pemilihan fitur ambang batas. Jumlah fitur yang tidak relevan dapat dikurangi dengan menghilangkan kata-kata berbobot rendah, sehingga meningkatkan efisiensi model dan menghasilkan performa klasifikasi yang konsisten di semua kasus uji.

5.2 Saran

1. Disarankan untuk penelitian berikutnya dapat mengembangkan kamus normalisasi bahasa informal atau bahasa media sosial yang digunakan pada tahap preprocessing. Sehingga proses normalisasi kata menjadi lebih akurat.
2. Penelitian selanjutnya dapat menyesuaikan pada pengaturan (parameter), seperti pembagian data latih dan data uji serta nilai threshold pada pembobotan

TF-IDF, untuk mengetahui pengaruh perubahan parameter tersebut terhadap performa model klasifikasi sentimen.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA

- Aditya Nugroho, R., & Cholissodin, I. (2021). *Implementasi Naïve Bayes Classifier untuk Klasifikasi Emosi Tweet Berbahasa Indonesia pada Spark* (Vol. 5, Issue 1).
- Aida, A. N., Arsi, P., Aji, R. P., & Tarwoto, T. (2024). *Analisis Sentimen Pengguna Aplikasi Instagram Pada Situs Google Play Menggunakan Metode Naïve Bayes*. *Jurnal Media Informatika Budidarma*, 8(2), 704. <https://doi.org/10.30865/mib.v8i2.7388>
- Anggi Pranata. (2022). *Penerapan Algoritma Support Vector Machine Klasifikasi Sentimen Analisis terhadap Opini Postingan Twitter (Maxim)*. *Jurnal Nasional Komputasi dan Teknologi Informasi* (Vol. 5, Issue 3)
- Fanissa, S., Fauzi, M. A., & Adinugroho, S. (2018). *Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking* (Vol. 2, Issue 8).
- Farah Zhafira, D., Rahayudi, B., & indriati. (2021). *Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF berdasarkan Komentar Pada Youtube* (Vol. 2, Issue 1).
- Fitrana, L. A., Linawati, S., & Herlinawati, N. (2024). *Analisis Sentimen Pengguna Twitter Terhadap Brand Indosat Menggunakan Metode Naive Bayes Classifier*. *JATI (Jurnal Mahasiswa Teknik Informatika)* (Vol. 8, Issue 3).
- Hasanah, A. N., & Sari, B. N. (2024). *Analisis Sentimen Ulasan Pengguna Aplikasi Jasa Ojek Online Maxim Pada Google Play Dengan Metode Naive Bayes Classifier*. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(1). <https://doi.org/10.23960/jitet.v12i1.3628>
- Joesyiana, K. (2018). *Pengaruh Word Of Mouth terhadap Keputusan Pembelian Konsumen pada Media Online Shop Shopee di Pekanbaru*. *Jurnal Valuta*, 4(1).
- Juniansih, S., Ripanti, E. F., & Pratama, E. E. (2020). *Implementasi Naive Bayes Classifier pada Opinion Mining Berdasarkan Tweets Masyarakat Terkait Kinerja Presiden dalam Aspek Ekonomi*. *Jurnal Sistem Dan Teknologi Informasi (Justin)*, 8(3), 239. <https://doi.org/10.26418/justin.v8i3.39118>
- Larasati, N. Y. (2025). *Analisis Sentimen Pada Komentar Aplikasi Sehat*

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Indonesiaku menggunakan Metode Machine Learning dan Klasifikasi Positif-Negatif. *Jurnal Ilmiah Matematika*, 13.

Melina Salsabila, S., Alim Murtopo, A., & Fadhilah, N. (2022). Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier. *Jurnal Minfo Polgan, Volume 11*. www.tokopedia.com

Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 5, Issue 2).

Prabowo, W. A., & Wiguna, C. (2021). Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM. *Jurnal Media Informatika Budidarma*, 5(1), 149. <https://doi.org/10.30865/mib.v5i1.2604>

Putra, M. W. A., Susanti, Erlin, & Herwin. (2020). Analisis Sentimen Dompot Elektronik Pada Twitter Menggunakan Metode Naïve Bayes Classifier. *IT Journal Research and Development*, 5(1), 72–86. [https://doi.org/10.25299/itjrd.2020.vol5\(1\).5159](https://doi.org/10.25299/itjrd.2020.vol5(1).5159)

Rahel Lina Simanjuntak, Theresia Romauli Siagian, Vina Anggriani, & Arnita Arnita. (2023). Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Shopee Dengan Menggunakan Algoritma Naïve Bayes. *Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 3(3), 23–39. <https://doi.org/10.55606/teknik.v3i3.2411>

Rizki Rinaldi, A., Perjuangan No, J., & Majasem Kec Kesambi Kota Cirebon, B. (2024). Penerapan Metode Naive Bayes Classifier Pada Analisis Sentimen Aplikasi Gopay. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 1).

Sari1, D. N., Adelia2, F., Rosdiana3, F., Butar4, B. B., & Hariyanto5, M. (2020). Analisa Sentimen Terhadap Review Produk Kecantikan Menggunakan Metode Naive Bayes classifier. In *JIKA: Vol. ISSN*.

Setiawan, R. P., Irawan, B., & Prihartono, W. P. (2025). Analisis Sentimen Ulasan Growtopia di Google Playstore Menggunakan Naive Bayes Classifier untuk Identifikasi Kebutuhan Pengguna. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(2). <https://doi.org/10.23960/jitet.v13i2.6415>

Setiawenda Putri, K., Setiawan, R., & Pambudi, A. (2023). Analisis Sentimen terhadap Brand Skincare Lokal Menggunakan Naive Bayes Classifier. In *Technologia*



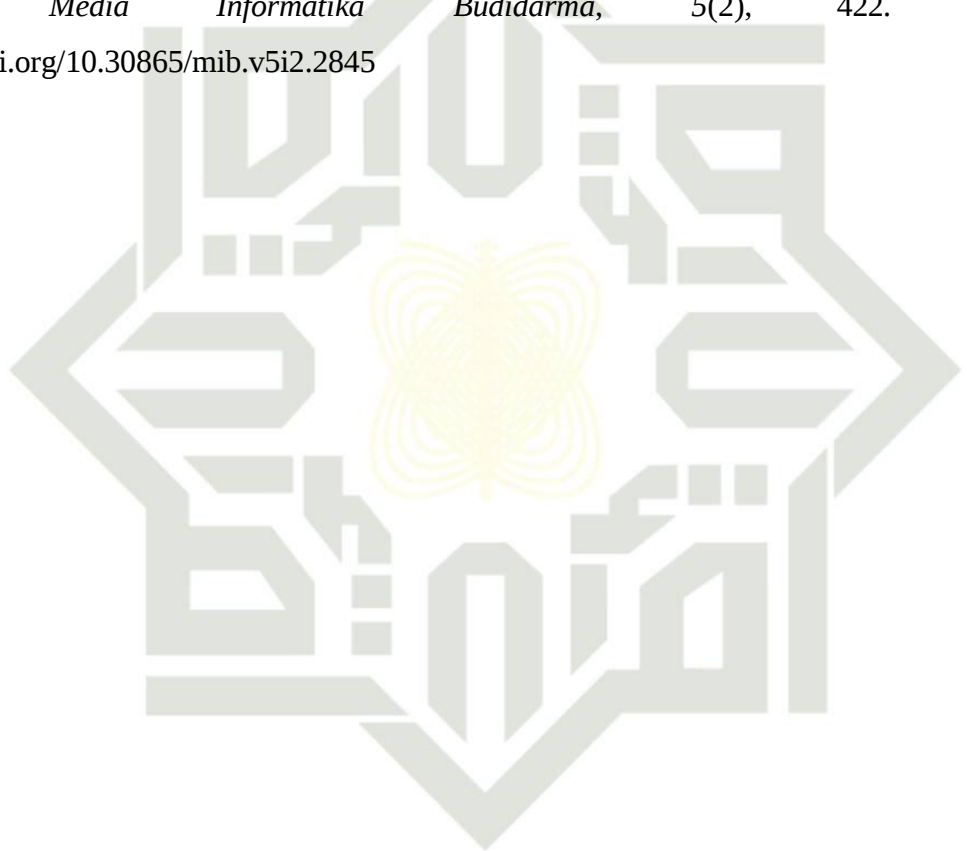
Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

(Vol. 14, Issue 3).

Wibowo, A. R., Astuti, Y. P., Kartikadarma, E., Subhiyakto, E. R., Anisa, N., Winarsih, S., & Rohman, M. S. (2024). *LogicLink : Journal of Artificial Intelligence and Multimedia in Informatics Penerapan Metode Naïve Bayes Classifier untuk Klasifikasi Sentimen pada Judul Berita* (Vol. 1, Issue 1).

Yutika, C. H., Adiwijaya, A., & Faraby, S. Al. (2021). Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes. *Jurnal Media Informatika Budidarma*, 5(2), 422. <https://doi.org/10.30865/mib.v5i2.2845>



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN A

	Username	At	Review	Kelas
3	Kenza Putra	2024-10-07 15:06:19	Aplikasi yg sangat mengecewakan. Kalo gk niat kasih vocer mending gk usah sekalian munculin vocer .. giliran digunakan langsung dibatalin sistem Mulu.. bikin kecewa gak kayak toko sbalah aman ² saja..	Negatif
	Hayyindiroh	2024-10-07 13:59:26	Masa setiap beli harga promo di batalin Mulu alesan nya punya data yang sama ga jelas banget, kalau mau bikin diskon yang benar dong jangan kayak gitu baru juga make dah kecewa	Negatif
	Egol Lawse	2024-10-07 11:03:15	Akun baru padahal tapi pas cekout dan promo udh dipake, tiba-tiba pembatalan sistem, kan aneh, jadi hangus promonya percuma download...gak jelas	Negatif
	Rangga Agung Nugraha	2024-10-07 07:59:28	Kalo sekiranya gak niat ngasih promo mending gak usah, ini saya udah klaim voucher tapi ketika digunakan malah dibatalkan otomatis oleh sistem, itu maksudnya apa? Ketika dicek lagi vouchernya tiba2 hilang. Emang gak jelas tokopedia.	Negatif
	Chrisna Paringah	2024-10-07 05:54:58	Dapat gratis ongkir tapi pengiriman nya lama sekali (harusnya 3 hari , ini sudah seminggu belum sampai) apalagi pasti pakai JNE.. Kalau bisa kontrak nya sama JNE di cut saja sangat mengecewakan pelanggan padahal barang nya sangat dibutuhkan..	Negatif



Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hendzky Vebrino	2024-10-07 05:48:43	Sekarang males pake tokopedia, akun saya sudah member platinum, tiba2 kupon gk bisa di gunakan, alasannya tidak sesuai syarat dan ketentuan, padahal saya tidak melanggar apapun, sangat kecewa.. lebih baik pindah ke shoope ja jauh lebih fair	Negatif
Dwi Susanto	2024-10-07 04:13:06	Selama bertahun 2 belanja di toped ini yg paling mengecewakan, belanja 2 kali dgn pengiriman yg 1 hari nyampek kenyataan sudah beberapa hari barang tdk kunjung datang, terus komplek harus menunggu lagi lama. Misalkan jasa kirim tdk memenuhi standar ndk usah dipakai lagi. Contoh aja aplikasi sebelah jasa kirim nya sedikit tapi profesional, Ini sudah teriak2 masih jg suruh nunggu . Kapokkkkkk blnja disini	Negatif
Sugiati	2024-10-07 00:05:09	saya pengguna baru toped,, saya kecewa dengan toped banyak pilihan produk promo pengguna baru ,tapi nggak bisa dibeli, misalkan bisa dibeli akan dibatalkan otomatis oleh sistem, payahhhh... TOPED SEGERA PERBAIKI SISTEM !!!	Negatif
Aji Pamungkas	2024-10-06 17:18:40	Maf tak kasih bintang 1,karna apk sering keluar sendiri .dan ngelag padahal sinyal bagus.tolong di perbaiki	

Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

			lagi.	
	Sohib Farhan	2024-10-06 15:47:13	Pengiriman lama, jawaban di penilaian templet semua, bukan solusi malah ucapan terimakasih menggunakan kalimat templet dan gada pembenahan. Baru sekali belanja di sini, bukanya di kasih nyaman malah di kecewakan. KAPOK	Negatif

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Username	At	Review	Kelas
1	Risty Diana	2024-10-07 09:40:30	aplikasi yang ga pelit sering ngasih diskon dan promo terus juga ngasih rekomendasi bagus yang kurang itu kirim pesan ke pengguna pas itu pernah pakai gopay di suruh ketik sandi tapi sandi nya gak kirim selain itu bagus	Positif
2	Ade Pani Saputra	2024-10-06 07:19:59	Toko pedia adalah aplikasi yg asyik dan menarik bisa menarik uang yg ada d dompet kita bisa menarik saldo kita dan juga bisa menarik hati kita ... Terkadang ada juga hal yang membuat kita menarik keinginan karna selalu ada gangguan tokped terus untungnya bukan gangguan jiwa kalau gangguan jiwa kita repot	Positif
3	M Saiful Arif	2024-10-06 04:12:30	Hai Tokped.. Aplikasi nya sudah oke banget. Tapi kalo free ongkir / kurir rekomendasi selalu di kasih ke JNE yang super lelet. Atau cross kurir yang bikin luamaa. Dengan senang hati kami beralih ke shopee. Gak tau mau kritik kemana, maaf saya tulis disini.	Positif
4	Joyosentiko	2024-10-07 03:48:14	Sekarang Tokopedia untuk masalah pengiriman kurir rekomendasi sangat lama plng cepat seminggu bisa lebih baru sampai,berapa pesanan terakhir sangat lama sampainya.Tolong diperhatikan ubtuk masalah kurir	Positif
5	Sapiul Vivo	2024-10-05 11:56:14	Sayang nya gak ada gratis ongkir untuk semua toko Sudah daftar Aplus 6 bln ternyata gak bisa di pakai untuk semua toko hanya toko yg berkode Aplus itupun masih bayar ongkir hanya potongan aja Akhirnya hanya sy pakai untuk belanja di sekitar saya Kalo voucher cashback emang sering ada Kalau bisa gratis ongkir pasti Tokopedia akan maju	Positif

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

8	Audy Reinhard (Odik)	2024-10-06 05:22:49	Tolong pihak tokopedia menghentikan kerjasama dengan pihak JNE. Karena masalah keterlambatan yang luar biasa dalam proses pengiriman. Masak paket dari sby ke malang aja sampek 6hr belum juga datang???!! Saya yakin paket itu hilang...ekspedisi JNE sangat tidak kredibel...	Positif
	Rzuull	2024-10-04 13:05:18	untuk promo setiap bulan dan lain lain bagus sih, tapi jangan lah setiap ada promo bulanan harus update, ampe kapan begini terus dah. jangan di read aja ya tokped. untuk keseluruhan baguss poll tapi jangan update terus jugaa makan30gb juga lama lama	Positif
	ic olv	2024-10-04 07:25:16	sejauh ini bagus, waktu itu ada kendala barang gak sampe tapi status selesai, saya ajukan banding sampe 2 baru di acc□. 1 hal yang saya harapi, ketika pesanan instan tolong kami sebagai buyer sediakan channel untuk bisa berkomunikasi dengan driver/kurir layak nya buyer berkomunikasi dengan seller, bukan hanya telpon, seperti aplikasi oren.	Positif
	Siti Hajar	2024-10-03 09:21:58	Saya senang belanja di Tokopedia, dulu saya sering belanja di si oren sama si biru. Tapi sekarang saya beralih ke Tokopedia. Tapi kenapa ini kalau pengiriman dari ekspedisi sicepat lama banget, sampai2 paket saya di retur, padahal saya pesan paket NON COD. Baru pertama kali saya pesan paket NON COD bisa diretur sama ekspedisinya. Untuk Tokopedia tolong perbaiki ekspedisinya kalau mau bersaing sama ecommerce yang lain. Karena yang lain pengirimannya sudah cepat, sehari bisa sampai. Terima	Positif

			kasih	
10	Fitriherpani Herpani	2024-10-01 12:15:49	senang belanja di tokopedia gercep dan bnyk promo dan toko yg di twarkan nya pun toko2 yg berkualitas walaupun ga pernah dpt diskon flash sale □.. tapi yaaa dibanding olshop lain tokped the best lah	Positif

Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN B

SURAT PERNYATAAN PAKAR BAHASA

Saya yang bertanda tangan dibawah ini :

Nama : Elvina MPd
 Tempat Tanggal Lahir : Peranap. 02 Januari 1990
 Pekerjaan : Dosen Bahasa Indonesia
 Alamat : Perumahan Pranata Cluster Blok C4 Jl.
 Swakarya Tuah Karya, Tampan, Pekanbaru

Menyatakan dengan sesungguhnya bahwa benar sebagai pakar yang memvalidasi untuk pelabelan sentiment yang bersifat positif dan negatif dalam ulasan aplikasi Tokopedia di Google Play Store sebanyak 3000 data asli dalam Tugas Akhir dari :

Nama : Putri Apriyawati
 Nim : 12050127491
 Jurusan : Teknik Informatika
 Judul Tugas Akhir : Klasifikasi Sentimen Aplikasi Tokopedia dengan
 Menggunakan Metode Naïve Bayes Classifier
 berdasarkan Ulasan di Google Play

Demikian surat ini saya buat dengan sesungguhnya tanpa paksaan dari pihak manapun, atas perhatiannya saya ucapkan terimakasih.

Pekanbaru, 16 Desember 2025

Yang membuat pernyataan,

Elvina, S.Pd., M.Pd

UIN SUSKA RIAU