

PERBANDINGAN METODE *NAÏVE BAYES*, *SUPPORT VECTOR MACHINE*(SVM) DAN *LONG SHORT-TERM MEMORY*(LSTM) UNTUK ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI SAYURBOX DI GOOGLE PLAY STORE

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh:

EKA ABIDAH ARDELIA
NIM : 09021282126106

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA

2025

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN METODE NAIVE BAYES, SUPPORT VECTOR MACHINE(SVM) DAN LONG SHORT-TERM MEMORY(LSTM) UNTUK ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI SAYURBOX DI GOOGLE PLAY STORE

Sebagai salah satu syarat untuk penyelesaian studi di

Program Studi S1 Teknik Informatika

Oleh:

EKA ABIDAH ARDELIA

09021282126106

Pembimbing 1 : Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

Mengetahui

Ketua Jurusan Teknik Informatika



Hadipurnawan Satria, Ph.D
198004182020121001

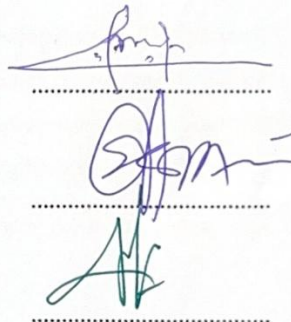
TANDA LULUS UJIAN KOMPREHENSIF

Pada hari jumat tanggal 13 Juni 2025 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Eka Abidah Ardelia
Nim : 09021282126106
Judul : Perbandingan Metode *Naive Bayes*, *Support Vector Machine*(SVM) dan *Long Short-Term Memory*(LSTM) Untuk Analisis Sentimen Ulasan Pengguna Aplikasi Sayurbox di Google Play Store

dan dinyatakan **LULUS**.

1. Ketua Penguji
Prof. Dr. Ermatita, M.Kom
NIP. 196709132006042001
2. Penguji I
Endang Lestari, S.Kom, M.T
NIP. 197811172006042001
3. Pembimbing I
Alvi Syahrini Utami, M.Kom
NIP. 197812222006042003



Mengetahui
Ketua Jurusan Teknik Informatika

Hadipurnawan Satria, Ph.D.
NIP. 198004182020121001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Eka Abidah Ardelia

NIM : 09021282126106

Program Studi : Teknik Informatika

Judul Skripsi : Perbandingan Metode *Naïve Bayes*, *Support Vector Machine*(SVM) dan *Long Short-Term Memory*(LSTM) Untuk Analisis Sentimen Ulasan Pengguna Aplikasi Sayurbox Di Google Play Store

Hasil Pengecekan *Software iThenticate/Turnitin* : 9%

Menyatakan bahwa Laporan Proyek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, 12 Juni 2025



Eka Abidah Ardelia
NIM. 09021282126106

MOTTO DAN PERSEMBAHAN

“Sesungguhnya beserta kesulitan ada kemudahan”

(Q.S Al-Insyirah : 6)

Kupersembahkan karya tulis ini kepada :

- Orang Tua, Saudara dan Keluargaku
- Diri saya sendiri
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

ABSTRACT

Sayurbox is an e-commerce application that provides fresh products directly from farmers to consumers. To understand user perception namely how users express their evaluations of the services and experiences when using the application this study conducts sentiment analysis on user reviews of Sayurbox from the Google Play Store using three classification algorithms: Naïve Bayes, Support Vector Machine (SVM), and Long Short-Term Memory (LSTM). Feature extraction is performed using TF-IDF for Naïve Bayes and SVM, and pre-trained Word2Vec for LSTM. The dataset consists of thousands of user reviews written in Indonesian. Evaluation results show that the SVM model with a linear kernel and $C=100$ achieves the best accuracy at 91.42%, followed by LSTM with 85.92%, and Naïve Bayes with 81.63%. These findings indicate that SVM is the most effective method for classifying the sentiment of Sayurbox user reviews.

Keywords: *Sentiment Analysis, Naïve Bayes, Support Vector Machine, LSTM, TF-IDF, Word2Vec, Sayurbox.*

ABSTRAK

Sayurbox merupakan aplikasi *e-commerce* yang menyediakan produk segar langsung dari petani ke konsumen. Untuk memahami persepsi pengguna, yaitu bagaimana pengguna menyampaikan penilaian terhadap layanan dan pengalaman menggunakan aplikasi, penelitian ini melakukan analisis sentimen terhadap ulasan pengguna Sayurbox di Google Play Store menggunakan tiga algoritma klasifikasi: *Naïve Bayes*, *Support Vector Machine* (SVM), dan *Long Short-Term Memory* (LSTM). Ekstraksi fitur dilakukan menggunakan TF-IDF untuk *Naïve Bayes* dan SVM, serta pre-trained Word2Vec untuk LSTM. Dataset yang digunakan terdiri dari ribuan ulasan berbahasa Indonesia. Hasil evaluasi menunjukkan bahwa model SVM dengan kernel linear dan parameter $C=100$ memberikan akurasi terbaik sebesar 91,42%, diikuti LSTM dengan akurasi 85,92% dan *Naïve Bayes* sebesar 81,63%. Temuan ini menunjukkan bahwa SVM merupakan metode paling efektif untuk klasifikasi sentimen ulasan pengguna Sayurbox.

Kata Kunci: Analisis Sentimen, *Naïve Bayes*, *Support Vector Machine*, LSTM, TF-IDF, *Word2Vec*, Sayurbox.

KATA PENGANTAR

Puji dan Syukur kehadirat Allah SWT atas berkat, rahmat dan karunia-Nya penulis dapat menyelesaikan tugas akhir yang berjudul “Perbandingan Metode *Naïve Bayes*, *Support Vector Machine*(SVM) Dan *Long Short-Term Memory*(LSTM) Untuk Analisis Sentimen Ulasan Pengguna Aplikasi Sayurbox Di Google Play Store” ini dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat lulus tingkat Strata-1 Program Studi Teknik Informatika pada Fakultas Ilmu Komputer Universitas Sriwijaya.

Penulis juga ingin mengucapkan banyak terima kasih kepada pihak-pihak yang telah memberikan dukungan, motivasi, bimbingan dan bantuan dalam menyelesaikan tugas akhir ini, yaitu:

1. Kepada kedua orang tuaku tersayang, Ahmad Syarifman, S.E dan Isabela, M.Kes. Terima kasih untuk perjuangan, doa yang tidak pernah putus, kasih sayang, dukungan, dan motivasi sehingga penulis bisa menyelesaikan skripsi ini untuk meraih gelar sarjana.
2. Kepada ketiga adikku tersayang, Caca, Hafiz, Salsa yang selalu mendoakan dan mendukung penulis.
3. Prof. Dr. Erwin, S.Si., M.Si selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Hadipurnawan Satria, S.Kom, M.Sc, Ph.D selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Alvi Syahrini Utami, M.Kom selaku Dosen Pembimbing yang telah membimbing dan memberikan masukan dalam proses pengerjaan tugas akhir.
6. Bapak Alfarissi, M.COMP.SC selaku Dosen Pembimbing Akademik di Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
7. Seluruh Bapak dan Ibu Dosen Program Studi Teknik Informatika yang telah memberikan ilmu kepada penulis selama masa perkuliahan.

8. Seluruh Staff Tata Usaha Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
9. M.Putra Apriansyah yang telah menemani, membantu, meluangkan waktunya, tenaga dan pikirannya, serta memberikan dukungan dan motivasinya selama ini. Semoga segala harapan baik yang telah direncanakan bisa terwujud dikemudian hari.
10. Kepada Nasywa, Dhifa, Athia, Laras dan Putri yang telah menemani penulis dari bangku sekolah menengah pertama, terima kasih untuk dukungan dan selalu ada suka maupun duka.
11. Kepada Msy Rizkia Yuniandari, teman seperjuangan selama masa perkuliahan sampai menyelesaikan tugas akhir.
12. Pihak-pihak lain yang telah memotivasi penulis dan memberikan dukungan namun tidak dapat disebutkan satu-persatu.

Penulis menyadari bahwa dalam penyusunan tugas akhir ini masih terdapat banyak kesalahan dan kekurangan yang disebabkan keterbatasan pengetahuan dan pengalaman penulis. Oleh karena itu, penulis mengharapkan kritik dan saran yang bersifat membangun agar dapat membantu penelitian selanjutnya. Akhir kata, semoga tugas akhir ini dapat memberikan manfaat bagi banyak orang.

Indralaya, 12 Juni 2025



Eka Abidah Ardelia
09021282126106

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	Kesalahan! Bookmark tidak ditentukan.
TANDA LULUS UJIAN KOMPREHENSIF ..	Kesalahan! Bookmark tidak ditentukan.
HALAMAN PERNYATAAN	Kesalahan! Bookmark tidak ditentukan.
MOTTO DAN PERSEMBAHAN	v
<i>ABSTRACT</i>	vi
ABSTRAK.....	viii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xvi
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan.....	I-1
1.2 Latar Belakang.....	I-1
1.3 Rumusan Masalah.....	I-4
1.4 Tujuan Penelitian	I-4
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah	I-5
1.7 Sistematika Penulisan	I-5
1.7.1 BAB I. PENDAHULUAN	I-6
1.7.2 BAB II. KAJIAN LITERATUR.....	I-6
1.7.3 BAB III. METODOLOGI PENELITIAN.....	I-6
1.7.4 BAB IV. PENGEMBANGAN PERANGKAT LUNAK.....	I-6
1.7.5 BAB V. HASIL DAN ANALISIS PENELITIAN.....	I-6
1.7.6 BAB VI. KESIMPULAN DAN SARAN	I-6
1.8 Kesimpulan	I-7
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan.....	II-1
2.2 Landasan Teori.....	II-1
2.2.1 Analisis Sentimen	II-1
2.2.2 Sayurbox	II-1
2.2.3 <i>Web Scraping</i>	II-2

2.2.4	<i>Natural Language Processing (NLP)</i>	II-3
2.2.5	Pra-Pengolahan Teks	II-4
2.2.6	<i>Undersampling</i>	II-6
2.2.7	<i>Multinomial Naïve Bayes</i>	II-7
2.2.8	<i>Support Vector Machine (SVM)</i>	II-8
2.2.9	<i>Long Short-Term Memory (LSTM)</i>	II-9
2.2.10	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	II-13
2.2.11	<i>Confusion Matrix</i>	II-15
2.2.12	<i>Rational Unified Process</i>	II-17
2.2.13	<i>Word Embedding</i>	II-18
2.2.14	<i>Word2Vec</i>	II-19
2.2.15	<i>Labelling Data</i>	II-20
2.3	Penelitian Lain Yang Relevan	II-20
2.4	Kesimpulan	II-22
BAB III METODOLOGI PENELITIAN.....		III-1
3.1	Pendahuluan.....	III-1
3.2	Pengumpulan Data	III-1
3.2.1	Jenis dan Sumber Data	III-1
3.2.2	Metode Pengumpulan Data.....	III-2
3.3	Tahapan Penelitian	III-3
3.3.1	Menentukan Kerangka Kerja Penelitian	III-4
3.3.2	Menentukan Kriteria Pengujian	III-16
3.3.3	Menentukan Format Data Pengujian.....	III-17
3.3.4	Menentukan Alat Bantu Penelitian.....	III-18
3.3.5	Melakukan Pengujian Penelitian.....	III-18
3.3.6	Melakukan Analisis Hasil Pengujian dan Membuat Kesimpulan Penelitian	III-19
3.4	Metode Pengembangan Perangkat Lunak	III-19
3.4.1	Fase Insepsi	III-20
3.4.2	Fase Elaborasi	III-20
3.4.3	Fase Konstruksi.....	III-20
3.4.4	Fase Transisi.....	III-21
3.5	Kesimpulan	III-21
BAB IV PENGEMBANGAN PERANGKAT LUNAK		IV-1
4.1	Pendahuluan	IV-1

4.2	Fase Insepsi.....	IV-1
4.2.1	Pemodelan Bisnis.....	IV-1
4.2.2	Kebutuhan Sistem	IV-3
4.2.3	Analisis dan Perancangan	IV-3
4.2.4	Implementasi.....	IV-28
4.3	Fase Elaborasi	IV-31
4.3.1	Pemodelan Bisnis.....	IV-31
4.3.2	Kebutuhan	IV-34
4.3.3	Analisis dan Perancangan	IV-34
4.4	Fase Konstruksi.....	IV-36
4.4.1	Kebutuhan	IV-37
4.4.2	Implementasi.....	IV-37
4.5	Fase Transisi.....	IV-41
4.5.1	Pemodelan Bisnis.....	IV-41
4.5.2	Kebutuhan	IV-41
4.5.3	Analisis dan Perancangan	IV-41
4.5.4	Implementasi.....	IV-42
4.6	Kesimpulan	IV-43
BAB V	HASIL DAN ANALISIS PENELITIAN	V-1
5.1	Pendahuluan.....	V-1
5.2	Data Hasil Penelitian.....	V-1
5.2.1	Konfigurasi Percobaan.....	V-1
5.2.2	Hasil Konfigurasi Model Naïve Bayes Skenario 1 (Imbalanced Dataset).....	V-5
5.2.3	Hasil Konfigurasi Model SVM Skenario 1 (Imbalanced Dataset).....	V-7
5.2.4	Hasil Konfigurasi Model LSTM Skenario 1 (Imbalanced Dataset).....	V-9
5.2.5	Hasil Konfigurasi Model Naïve Bayes Skenario 2 (Balanced Dataset)	V-11
5.2.6	Hasil Konfigurasi Model SVM Skenario 2 (Balanced Dataset)	V-14
5.2.7	Hasil Konfigurasi Model LSTM Skenario 2 (Balanced Dataset)	V-16
5.3	Analisis Hasil Penelitian	V-18
5.4	Kesimpulan	V-20
BAB VI	KESIMPULAN DAN SARAN	VI-1
6.1	Pendahuluan.....	VI-1
6.2	Kesimpulan	VI-1

6.3	Saran	VI-2
DAFTAR PUSTAKA.....		xviii
LAMPIRAN.....		xxii

DAFTAR TABEL

Tabel II- 1. Confusion Matrix 2x2	II-16
Tabel III- 1. Contoh data review yang dikumpulkan	III-1
Tabel III- 2. Pembagian Data Positif dan Data Negatif	III-11
Tabel III- 3. Pembagian Data Latih dan Data Uji	III-11
Tabel III- 4. Pembagian Data Setelah Undersampling.....	III-12
Tabel III- 5. Distribusi Sentimen pada Data Latih dan Data Uji (Setelah Undersampling)	III-13
Tabel III- 6. Rancangan Tabel Pengujian Confusion Matrix	III-17
Tabel III- 7. Rancangan Tabel Hasil Pengujian	III-17
Tabel IV- 1. Kebutuhan Fungsional	IV-3
Tabel IV- 2. Kebutuhan Non-Fungsional	IV-3
Tabel IV- 3. Contoh Data Ulasan Pengguna Sayurbox	IV-6
Tabel IV- 4. Hasil Proses Hapus Data Duplikat dan Minim Informasi	IV-7
Tabel IV- 5. Hasil Proses Cleaning	IV-7
Tabel IV- 6. Hasil Proses Case Folding.....	IV-8
Tabel IV- 7. Hasil Proses Normalization.....	IV-9
Tabel IV- 8. Hasil Proses Tokenizing	IV-10
Tabel IV- 9. Hasil Proses Stopword Removal.....	IV-11
Tabel IV- 10. Hasil Proses Stemming	IV-11
Tabel IV- 11. Hasil Perhitungan TF dan IDF	IV-13
Tabel IV- 12. Hasil Perhitungan TF-IDF.....	IV-13
Tabel IV- 13. Token Angka	IV-15
Tabel IV- 14. Hasil Padding	IV-15
Tabel IV- 15. Hasil Word2Vec.....	IV-17
Tabel IV- 16. Definisi Aktor.....	IV-28
Tabel IV- 17. Definisi Use Case	IV-29
Tabel IV- 18. Skenario Use Case Melakukan Analisis Sentimen.....	IV-29
Tabel IV- 19. Implementasi Kelas Berdasarkan Class Diagram	IV-37
Tabel IV- 20. Daftar File Python.....	IV-38
Tabel IV- 21. Rancangan Pengujian Use Case Input Kalimat Ulasan.....	IV-40
Tabel IV- 22. Rancangan Pengujian Use Case Pemilihan Model	IV-40

Tabel IV- 23.	Rancangan Pengujian Use Case Analisis Sentimen	IV-41
Tabel IV- 24.	Pengujian Use Case Input Kalimat Ulasan	IV-41
Tabel IV- 25.	Pengujian Use Case Pemilihan Model	IV-41
Tabel IV- 26.	Pengujian Use Case Analisis Sentimen.....	IV-42
Tabel V- 1.	Konfigurasi Parameter Naive Bayes.....	V-2
Tabel V- 2.	Konfigurasi Parameter SVM	V-2
Tabel V- 3.	Konfigurasi Parameter LSTM	V-2
Tabel V- 4.	Konfigurasi Model Naive Bayes.....	V-3
Tabel V- 5.	Konfigurasi Model SVM	V-4
Tabel V- 6.	Konfigurasi Model LSTM	V-4
Tabel V- 7.	Hasil Pelatihan Model Naive Bayes dengan Imbalanced Dataset	V-5
Tabel V- 8.	Hasil Pelatihan Model SVM dengan Imbalanced Dataset	V-7
Tabel V- 9.	Hasil Pelatihan Model LSTM dengan Imbalanced Dataset	V-10
Tabel V- 10.	Hasil Pelatihan Naive Bayes dengan Balanced Dataset	V-12
Tabel V- 11.	Hasil Pelatihan Model SVM dengan Balanced Dataset	V-14
Tabel V- 12.	Hasil Pelatihan Model LSTM dengan Balanced Dataset.....	V-16

DAFTAR GAMBAR

Gambar II- 1. Arsitektur Long Short-Term Memory	II-11
Gambar III- 1. Tahapan Penelitian.....	III-4
Gambar III- 2. Diagram Alur Kerangka Kerja Penelitian.....	III-5
Gambar III- 3. Flowchart Preprocessing.....	III-9
Gambar IV- 1. Use Case Diagram.....	IV-28
Gambar IV- 2. Rancangan Antarmuka Sistem Analisis Sentimen.....	IV-32
Gambar IV- 3. Rancangan Antarmuka Dropdown Pilihan Model Terbuka.....	IV-32
Gambar IV- 4. Rancangan Antarmuka Hasil Analisis Sentimen	IV-33
Gambar IV- 5. Activity Diagram.....	IV-34
Gambar IV- 6. Sequence Diagram	IV-35
Gambar IV- 7. Class Diagram Perangkat Lunak	IV-36
Gambar IV- 8. Antarmuka Sistem Ulasan Pengguna	IV-39
Gambar IV- 9. Antarmuka Dropdown Pilihan Model Terbuka	IV-39
Gambar IV- 10. Antarmuka Hasil Analisis Sentimen	IV-39
Gambar V- 1. Confusion Matrix Hasil Naive Bayes Imbalanced Dataset.....	V-6
Gambar V- 2. Confusion Matrix Hasil SVM Imbalanced Dataset	V-8
Gambar V- 3. Confusion Matrix Hasil LSTM Imbalanced Dataset	V-11
Gambar V- 4. Confusion Matrix Hasil Naive Bayes Balanced Dataset	V-13
Gambar V- 5. Confusion Matrix Hasil SVM Balanced Dataset	V-15
Gambar V- 6. Confusion Matrix Hasil LSTM Balanced Dataset	V-17

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab ini menyajikan uraian mengenai dasar pemikiran munculnya permasalahan, perumusan masalah yang akan dikaji, tujuan dari penelitian yang dilakukan, kegunaan hasil penelitian, ruang lingkup permasalahan, struktur penulisan laporan, serta rangkuman dari topik tugas akhir yang diangkat.

1.2 Latar Belakang

Perkembangan teknologi digital mengalami akselerasi yang signifikan dalam dekade terakhir, memberikan dampak transformatif pada berbagai aspek kehidupan masyarakat. Transformasi digital ini telah mengubah pola konsumsi dan belanja masyarakat Indonesia dari cara konvensional menjadi berbasis aplikasi (Kurniawan et al., 2022). Perkembangan teknologi digital semakin mempercepat transformasi cara masyarakat dalam memenuhi kebutuhan pokok sehari-hari, terutama dalam hal pengadaan bahan pangan.

Sejalan dengan perkembangan tersebut, Sayurbox hadir sebagai platform *e-commerce* yang berfokus pada penyediaan sayuran dan bahan makanan segar langsung dari petani ke konsumen. Dibentuk pada tahun 2016, aplikasi ini telah berkembang pesat dengan menghadirkan lebih dari 5.000 produk segar dan melayani berbagai kota besar di Indonesia (Firdaus et al., 2023). Data dari Google Play Store menunjukkan Sayurbox telah diunduh lebih dari 1 juta kali dengan rating

4,2 dan ulasan mencapai 98 ribu. Fitur ulasan ini menjadi sumber informasi berharga bagi calon pengguna dan pengembang aplikasi untuk memahami kualitas layanan yang diberikan (Wijaya & Santoso, 2022). Meskipun terdapat berbagai ulasan positif yang menonjolkan kualitas produk, terdapat pula beragam tanggapan kritis seputar aspek pengiriman dan harga. Guna memperoleh pemahaman mendalam tentang dinamika sentimen pengguna, penelitian ini menggunakan metode *text mining* untuk melakukan analisis secara menyeluruh dan objektif.

Analisis sentimen merupakan cabang penting dari *text mining* yang berfokus pada identifikasi dan ekstraksi opini dari teks tidak terstruktur menggunakan kombinasi pemrosesan bahasa alami dan algoritma pembelajaran mesin (Purwanto & Djamal, 2023). Implementasi analisis sentimen pada ulasan aplikasi memberikan wawasan mendalam tentang persepsi pengguna, sehingga pengembang dapat mengidentifikasi area yang perlu ditingkatkan dan mempertahankan fitur yang disukai (Azizah & Windarti, 2023).

Dalam penelitian ini, tiga algoritma pembelajaran mesin dipilih untuk dibandingkan performanya dalam klasifikasi sentimen: *Naïve Bayes*, *Support Vector Machine* (SVM), dan *Long Short-Term Memory* (LSTM). Pemilihan *Naïve Bayes* didasarkan pada kesederhanaannya, kecepatan komputasi yang tinggi, serta kemampuannya dalam menangani data berfitur besar meskipun ukuran dataset terbatas. SVM dipilih karena kekuatannya dalam menangani data berdimensi tinggi serta kestabilannya dalam berbagai konteks analisis sentimen, terutama untuk data teks berformat TF-IDF. Adapun LSTM, yang merupakan bagian dari pendekatan *deep learning*, digunakan karena kemampuannya memahami konteks dan pola

urutan kata dalam teks, sehingga sangat sesuai untuk mendeteksi opini yang tersirat atau kompleks. Ketiga algoritma ini mencerminkan pendekatan berbeda yang saling melengkapi, mulai dari klasifikasi probabilistik sederhana, pemodelan margin optimal, hingga pemrosesan urutan berbasis konteks mendalam (Kristiyanti & Hardani, 2023; Nikmah et al., 2022).

Penelitian terdahulu oleh Nikmah et al. (2022) menunjukkan bahwa metode SVM memberikan akurasi tertinggi (86,54%) dalam mengklasifikasi tweet pelecehan seksual berbahasa Indonesia dibandingkan LSTM (84,62%) dan *Naïve Bayes* (85,45%). Sementara itu, Kristiyanti dan Hardani (2023) melakukan studi analisis sentimen terhadap ulasan vaksin Covid-19 di Indonesia dan melaporkan bahwa SVM kembali unggul dengan akurasi 84,89%, disusul oleh *Naïve Bayes* (84,65%) dan LSTM (82,97%). Kedua studi tersebut mendemonstrasikan bahwa meskipun ketiganya menunjukkan performa baik, terdapat variasi efektivitas tergantung pada konteks dan karakteristik data.

Berdasarkan latar belakang tersebut, tujuan penelitian ini adalah melakukan analisis sentimen terhadap ulasan pengguna aplikasi Sayurbox berbahasa Indonesia di Google Play Store dan membandingkan performa tiga algoritma *machine learning*: *Naive Bayes*, *Support Vector Machine* (SVM), dan *Long Short-Term Memory* (LSTM). Penelitian ini diharapkan dapat menghasilkan temuan signifikan tentang algoritma mana yang paling efektif dalam konteks analisis sentimen ulasan aplikasi *e-commerce* pangan, sehingga dapat memberikan kontribusi praktis bagi pengembangan sistem analisis sentimen yang lebih optimal. Selain itu, hasil analisis sentimen dapat menjadi masukan strategis bagi tim pengembang Sayurbox untuk

meningkatkan kualitas layanan dan pengalaman pengguna berdasarkan umpan balik yang diberikan.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, adapun rumusan masalah pada penelitian ini yaitu sebagai berikut :

1. Bagaimana implementasi analisis sentimen terhadap ulasan pengguna aplikasi Sayurbox di Google Play Store menggunakan algoritma *Naïve Bayes*, *Support Vector Machine* dan *Long Short-Term Memory*?
2. Bagaimana performa dari algoritma *Naïve Bayes*, *Support Vector Machine*, dan *Long Short-Term Memory* dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Sayurbox di Google Play Store?

1.4 Tujuan Penelitian

Adapun tujuan penelitian ini yaitu sebagai berikut :

1. Mengimplementasikan analisis sentimen terhadap ulasan pengguna aplikasi Sayurbox di Google Play Store menggunakan algoritma *Naïve Bayes*, *Support Vector Machine*, dan *Long Short-Term Memory*.
2. Membandingkan performa tiga algoritma yaitu *Naive Bayes*, *Support Vector Machine* dan *Long Short-Term Memory* dalam menganalisis sentimen ulasan pengguna aplikasi Sayurbox di Google Play Store.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini yaitu sebagai berikut :

1. Mendapatkan hasil analisis sentimen ulasan pengguna aplikasi Sayurbox di Google Play Store.
2. Memberikan informasi mengenai perbandingan algoritma *Naive Bayes*, *Support Vector Machine* (SVM), dan *Long Short-Term Memory* (LSTM) dalam menganalisis sentimen ulasan pengguna aplikasi Sayurbox di Google Play Store.
3. Membantu tim pengembang Sayurbox dalam memahami persepsi dan pengalaman pengguna melalui interpretasi sentimen ulasan, yang dapat digunakan untuk meningkatkan kualitas dan fungsionalitas aplikasi.
4. Memberi gambaran kepada calon pengguna tentang kualitas dan kinerja aplikasi Sayurbox berdasarkan analisis sentimen ulasan pengguna.

1.6 Batasan Masalah

Batasan masalah dari penelitian ini yaitu sebagai berikut :

1. Data ulasan pengguna yang dianalisis adalah ulasan berbahasa Indonesia yang tersedia di platform Google Play Store.
2. Klasifikasi sentimen ulasan pengguna dibatasi pada dua kelas, yaitu sentimen positif dan sentimen negatif.
3. Evaluasi kinerja algoritma *Naive Bayes*, *Support Vector Machine* (SVM), dan *Long Short-Term Memory* (LSTM) dilakukan menggunakan *confusion matrix*, dengan parameter akurasi, presisi, *recall*, dan *F1-score*.

1.7 Sistematika Penulisan

Struktur penulisan tugas akhir ini disusun sebagai berikut:

1.7.1 BAB I. PENDAHULUAN

Bab ini menjelaskan secara lengkap mengenai latar belakang permasalahan, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, sistematika penulisan, serta ringkasan isi dari bab ini.

1.7.2 BAB II. KAJIAN LITERATUR

Bab ini menguraikan teori-teori dan literatur yang mendasari serta berhubungan dengan pelaksanaan penelitian ini.

1.7.3 BAB III. METODOLOGI PENELITIAN

Bab ini memaparkan tahapan-tahapan penelitian yang dilakukan. Setiap tahap dijelaskan secara terstruktur dalam bentuk kerangka kerja.

1.7.4 BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini menjelaskan proses pembangunan perangkat lunak berdasarkan pendekatan *Rational Unified Process* (RUP), termasuk proses pengujian terhadap perangkat lunak yang telah dibuat.

1.7.5 BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini memuat paparan detail mengenai hasil implementasi dan pengujian perangkat lunak, serta analisis terhadap hasil tersebut.

1.7.6 BAB VI. KESIMPULAN DAN SARAN

Bab ini merangkum hasil penelitian dalam bentuk kesimpulan serta memberikan saran yang dapat digunakan sebagai acuan untuk penelitian lebih lanjut.

1.8 Kesimpulan

Berdasarkan uraian yang telah disampaikan, penelitian ini akan dilaksanakan sesuai dengan metode yang telah ditentukan serta mengikuti sistematika penulisan ilmiah yang runtut dan terstruktur.

DAFTAR PUSTAKA

- Ahmad, M., et al. (2022). "Improving Word Vector Representations for Sentiment Analysis in Indonesian Language." *Journal of Computer Science*, 18(2), 45-56.
- Af'idah, D. I., Dairoh, & Sari, S. N. (2021). *Bidirectional Long Short Term Memory dan Word2Vec Untuk Analisis Sentimen Terhadap Ulasan Destinasi Wisata Pulau Bali*. http://eprints.poltektegal.ac.id/974/1/DwiIntanAf%27idah_11020470_penelitian.pdf
- Andi, A. W. (2020). *Analisis sentimen pemindahan Ibu Kota Indonesia menggunakan algoritma Deep Learning* [UIN Sunan Gunung Djati Bandung]. <https://etheses.uinsgd.ac.id/33950/>
- Atimi, R. L., & Enda Esyudha Pratama. (2022). Implementasi Model Klasifikasi Sentimen Pada Review Produk Lazada Indonesia. *Jurnal Sains Dan Informatika*, 8(1), 88–96. <https://doi.org/10.34128/jsi.v8i1.419>
- Baskara, R., & Rahma, F. (2022). Implementasi Web Scraping Pada Media Sosial Instagram. *Automata*, 3, 1–3.
- Ahmad, M., et al. (2022). "Improving Word Vector Representations for Sentiment Analysis in Indonesian Language." *Journal of Computer Science*, 18(2), 45-56.
- Hidayat, R., et al. (2023). "Performance Evaluation Metrics in Machine Learning: A Comprehensive Review." *International Journal of Artificial Intelligence*, 15(1), 112-125.
- Nugroho, A., et al. (2023). "LSTM-based Sentiment Analysis for Indonesian E-commerce Reviews." *Journal of Information Technology*, 6(2), 78-89.

- Pratama, B., et al. (2021). "Automated Data Collection Methods for E-commerce Analysis." *Digital Business Review*, 4(3), 167-178.
- Putri, S., et al. (2022). "Text Preprocessing Techniques for Indonesian Language Processing." *Journal of Data Science*, 8(4), 234-245.
- Rahman, F., et al. (2022). "Sentiment Analysis in E-commerce: Methods and Applications." *International Journal of Technology*, 13(2), 345-358.
- Susanto, Y., et al. (2023). "Advanced NLP Techniques for Indonesian Language Processing." *Asian Journal of Computer Science*, 12(1), 89-102.
- Wijaya, D., et al. (2023). "The Evolution of E-commerce Platforms in Southeast Asia." *Journal of Business Technology*, 10(1), 23-34.
- Susanti, A., Notowidigdo, R., & Trisnawati, M. (2020). *Farm-to-table supply chain and digital transformation: Case of Sayurbox in Indonesia*.
- Khorasani, M., Abdou, M., & Fernandez, J.H. (2022). *Web Application Development with Streamlit: Develop and Deploy Secure and Scalable Web Application to the Cloud Using a Pure PythonFramework*.
<https://www.researchgate.net/publication/362984327>
- Napoles, J.M., Biswas, A., Parker, M.I., & Rodríguez-Valdez, L. M. (2022). *Stmol: A component for building interactive molecular visualizations within streamlit web-applications*.
<https://www.researchget.net/publication/371327564>
- Ilmawan, L. B. (2018). Membangun Web Crawler Berbasis Web Service Untuk Data Crawling Pada Website Google Play Store. *ILKOM Jurnal Ilmiah*, 10(2), 215–224.
<https://doi.org/10.33096/ilkom.v10i2.282.215-224>

- Mustolih, A., Arsi, P., & Subarkah, P. (2023). Sentiment Analysis Motorku X using Applications Naive Bayes Classifier Method. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIMD)*, 6(2), 231 – 242. <https://doi.org/10.24014/ijaidm.v6i2.24864>
- Nufairi, F., Pratiwi, N., & Herlando, F. (2024). Analisis Sentimen Pada Ulasan Aplikasi Threads Di Google Play Store Menggunakan Algoritma Support Vector Machine. *JIPi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9(1), 339–348. <https://doi.org/10.29100/jipi.v9i1.4929>
- Nurdy, A. H., Rahim, A., & Arbansyah. (2024). Analisis Sentimen Ulasan Game Stumble Guys Pada Playstore Menggunakan Algoritma Naïve Bayes Sentiment Analysis of Stumble Guys Game Reviews on Playstore Using the Naïve Bayes Algorithm. *Teknika*, 13(November), 388–395. <https://doi.org/10.34148/teknika.v13i3.993>
- Nurian, A., & Nurina Sari, B. (2023). Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 11(3).
- Nurmalasari, D., Hermanto, T. I., & Nugroho, I. M. (2023). Perbandingan Algoritma SVM, KNN dan NBC Terhadap Analisis Sentimen Aplikasi Loan Service. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(3), 1521–1530.
- Permana, G. P., Nugraha, D. A., & Santoso, H. (2024). Perbandingan Performa SVM dan Naïve Bayes Pada Analisis Sentimen Aplikasi Game Online. *JOINTECS (Journal of Information Technology and Computer Science)*, 8(1), 21–30.
- Rifa'i, A., Sujaini, H., & Prawira, D. (2021). Sentiment Analysis Objek Wisata Kalimantan Barat Pada Google Maps Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 7(3), 400–407. <https://doi.org/10.26418/jp.v7i3.48132>

- Rohmansa, R. Q., Pratiwi, N., & Palepa, M. J. (2024). Analisis Sentimen Ulasan Pengguna Aplikasi Discord Menggunakan Metode K-Nearest Neighbor. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9(1), 368–378. <https://doi.org/10.29100/jipi.v9i1.4943>
- Romaito, E. S., Anam, M. K., Rahmaddeni, & Ulfah, A. N. (2021). Perbandingan Algoritma SVM Dan NBC Dalam Analisa Sentimen Pilkada Pada Twitter. *CSRIDJournal*, 13(3), 169–179. <https://doi.org/10.22303/csrid.13.3.2021.169-179>
- Santi, R. J., Setiawa, D., & Pratiwi, I. A. (2021). Perubahan Tingkah Laku Anak Sekolah Dasar Akibat Game Online. *Jurnal Penelitian Dan Pengembangan Pendidikan*, 5(3), 385–390. <https://doi.org/10.23887/jppp.v5i3.38576>
- Sapitri, I. A., Yusra, & Fikry, M. (2023). Pengklasifikasian Sentimen Ulasan Aplikasi Whatsapp Pada Google Play Store Menggunakan Support Vector Machine. *Jurnal Tekinkom (Teknik Informasi Dan Komputer)*, 6(1), 1–7. <https://doi.org/https://doi.org/10.37600/tekinkom.v6i1.773>
- Azrul, A., Irma Purnamasari, A., & Ali, I. (2024). Analisis Sentimen Pengguna Twitter Terhadap Perkembangan Artificial Intelligence Dengan Penerapan Algoritma Long Short-Term Memory (Lstm). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 413–421. <https://doi.org/10.36040/jati.v8i1.8416>

LAMPIRAN

Link Github : <https://github.com/EkaAbidahArdelia/Skripsi>