

**ANALISIS SENTIMEN ULASAN APLIKASI LEMON8 PADA
GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA
NAIVE BAYES**

TUGAS AKHIR

Disusun Oleh:

Edgard Sean Oley

22013011

UKDLSM



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS KATOLIK DE LA SALLE MANADO
MANADO
2026**

**ANALISIS SENTIMEN ULASAN APLIKASI LEMON8 PADA
GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA
NAIVE BAYES**

TUGAS AKHIR

Ditulis untuk memenuhi sebagian persyaratan untuk mendapatkan gelar Sarjana
Teknik pada Program Studi Teknik Informatika

Disusun Oleh:

Edgard Sean Oley

22013011



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS KATOLIK DE LA SALLE MANADO
MANADO
2026**

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Edgard Sean Oley
NIM : 22013011
Tempat/Tanggal Lahir : Manado, 22 Maret 2004
Fakultas/Program Studi : Teknik/Teknik Informatika

Menyatakan bahwa Laporan Tugas Akhir dengan judul "Analisis Sentimen Ulasan Aplikasi Lemon8 pada Google Play Store Menggunakan Algoritma Naive Bayes" yang penulis susun adalah benar hasil karya penulis dan bukan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian pernyataan ini penulis buat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar, maka penulis bersedia menerima sanksi akademis sesuai dengan yang diterapkan oleh Fakultas Teknik, berupa pembatalan Tugas Akhir.

Manado, 14 Juli 2026
Yang Membuat Pernyataan,


Edgard Sean Oley
Menyetujui,

Dosen Pembimbing I

Dosen Pembimbing II

Lanny Sitanayah, S.T., M.Sc., Ph.D.

Junaidy Budi Sanger, S.Kom., M.Kom

Mengetahui,

Ketua Program Studi

Dekan Fakultas Teknik

Vivie Deyby Khumenap, S.T., M.Cs.

Ronald Albert Rachmadi, S.T., M.T.



**UNIVERSITAS KATOLIK DE LA SALLE
MANADO-INDONESIA**

Nama : Edgard Sean Oley
NIM : 22013011
Fakultas : Teknik
Program Studi : Teknik Informatika
Judul Tugas Akhir : Analisis Sentimen Ulasan Aplikasi Lemon8 pada Google
Play Store Menggunakan Algoritma Naive Bayes
Pembimbing I : Lanny Sitanayah, S.T., M.Sc., Ph.D.
Pembimbing II : Junaidy Budi Sanger, S.Kom., M.Kom.

Manado, 14 Juli 2026
Menyetujui,

Dosen Pembimbing I
^

Dosen Pembimbing II

Lanny Sitanayah, S.T., M.Sc., Ph.D. Junaidy Budi Sanger, S.Kom., M.Kom.
Mengetahui,
Ketua Program Studi Dekan Fakultas Teknik

Vivie Deyby Kumenap, S.T., M.Cs.

Ronald Albert Rachmadi, S.T., M.T.

ABSTRACT

User reviews on the Lemon8 application contain crucial information regarding audience satisfaction, but the large volume of data complicates manual analysis. This study aims to build a Streamlit-based sentiment analysis software using the Multinomial Naive Bayes algorithm to classify reviews into Positive, Negative, and Neutral classes. Text data was processed using PySastrawi and extracted with the TF-IDF method. The computational evaluation was conducted through two testing scenarios using an optimal 80:20 data split ratio: Scenario 1 using the original imbalanced dataset (3,914 reviews) and Scenario 2 using a balanced dataset via undersampling (1,131 reviews). Functionality testing (Black Box) showed that all application features operated with 100% validity. In algorithmic performance testing, Scenario 1 achieved an average accuracy of 80.72% but suffered from majority class bias, failing to detect the Neutral class. Conversely, Scenario 2 yielded an average accuracy of 74.63% with a much more objective and proportional classification performance across classes. This research concludes that balancing data distribution is crucial in the sentiment classification process to prevent prediction bias toward the majority class.

Keywords: Sentiment Analysis, Lemon8, Multinomial Naive Bayes, Data Imbalance, Streamlit.

ABSTRAK

Ulasan pengguna pada aplikasi Lemon8 memuat informasi penting terkait tingkat kepuasan audiens, namun, besarnya volume data menyulitkan analisis secara manual. Penelitian ini bertujuan membangun aplikasi analisis sentimen berbasis Streamlit menggunakan algoritma Multinomial Naive Bayes guna mengklasifikasikan ulasan ke dalam kelas Positif, Negatif, dan Netral. Data teks diproses menggunakan PySastrawi dan diekstraksi dengan metode TF-IDF. Evaluasi komputasi dilakukan melalui dua skenario pengujian dengan rasio pembagian data optimal 80:20, yaitu Skenario 1 menggunakan *dataset* asli yang tidak seimbang (3.914 ulasan) dan Skenario 2 menggunakan *dataset* seimbang hasil *undersampling* (1.131 ulasan). Hasil pengujian fungsionalitas (*Black Box*) menunjukkan seluruh fitur aplikasi berjalan valid 100%. Pada pengujian performa algoritma, Skenario 1 mencatat akurasi rata-rata 80,72% namun mengalami *majority class bias* sehingga gagal mendeteksi kelas Netral. Sebaliknya, Skenario 2 menghasilkan akurasi rata-rata 74,63% dengan kinerja klasifikasi per kelas yang jauh lebih objektif dan proporsional. Penelitian ini menyimpulkan bahwa penyeimbangan distribusi data sangat krusial dalam proses klasifikasi sentimen untuk mencegah bias prediksi terhadap kelas mayoritas.

Kata Kunci: Analisis Sentimen, Lemon8, *Multinomial Naive Bayes*, Ketidakseimbangan Data, Streamlit.

KATA PENGANTAR

Diucapkan puji dan syukur serta terima kasih dipanjatkan kepada Tuhan Yesus Kristus karena dengan berkat kemurahan-Nya sehingga laporan dan pembuatan aplikasi Analisis Sentimen Ulasan Aplikasi Lemon8 pada Google Play Store Menggunakan Algoritma Naive Bayes untuk Tugas Akhir ini dapat berlangsung dengan lancar. Dalam penulisan laporan dan penyusunan laporan dan aplikasi tentunya terdapat beberapa pihak yang membantu sehingga boleh selesai dengan baik. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Dr. Gregorius Hertanto Dwi Wibowo, S.S., M.Th. selaku Rektor Universitas Katolik De La Salle Manado.
2. Bapak Ronald Albert Rachmadi, S.T., M.T. selaku Dekan Fakultas Teknik.
3. Ibu Vivie Deyby Kumenap, S.T., M.Cs. selaku Ketua Program Studi Teknik Informatika.
4. Ibu Lanny Sitanayah, S.T., M.Sc., Ph.D. selaku Dosen Pembimbing I yang selalu mengarahkan dan membantu sehingga penulis dapat menyelesaikan proses penyusunan laporan Tugas Akhir ini.
5. Bapak Junaidy Budi Sanger, S.Kom., M.Kom. selaku Dosen Pembimbing II yang juga telah memberikan arahan dalam penyusunan laporan Tugas Akhir ini.
6. Keluarga penulis yang selalu mendukung penulis dan doa yang terbaik.
7. Tetangga dan saudara di gereja yang selalu mengingatkan penulis untuk menyelesaikan laporan Tugas Akhir ini.
8. Miracle, Yoga, Hizkia, Christiano, Rafael, Gerald, Alfian, Jesco, Brooklyn, Steven, Jonas, Jasson, Dhiego, Marcelino, Vincent, Andrew, Algy, Immanuel, Harly, Rafelshen, Clay, Efraim, Brandon, Marvil dan Rafael M selaku sahabat saya yang mendukung saya selama menyusun laporan Tugas Akhir.
9. Teman-teman seperjuangan angkatan 2022 Fakultas Teknik yang selalu saling membantu dan menyemangati satu sama lain.

Manado, Juli 2026



Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PERNYATAAN.....	ii
LEMBAR PENGESAHAN	iii
ABSTRACT.....	iv
ABSTRAK	v
KATA PENGANTAR.....	vi
DAFTAR ISI	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
 BAB I PENDAHULUAN	 11
1.1 Latar Belakang	11
1.2 Rumusan Masalah	13
1.3 Tujuan Tugas Akhir	13
1.4 Manfaat Tugas Akhir.....	13
1.5 Batasan Masalah.....	13
1.6 Metodologi Penelitian	14
1.7 Sistematika Penulisan	15
 BAB II STUDI PUSTAKA	 Error! Bookmark not defined.
2.1 Teori Pendukung	Error! Bookmark not defined.
2.1.1 Analisis Sentimen	Error! Bookmark not defined.
2.1.2 Google Play Store	Error! Bookmark not defined.
2.1.3 Aplikasi Lemon8.....	Error! Bookmark not defined.
2.1.4 Web <i>Scraping</i>	Error! Bookmark not defined.
2.1.5 Preprocessing Teks.....	Error! Bookmark not defined.
2.1.6 TF-IDF (<i>Term Frequency-Inverse Document Frequency</i>).....	Error! Bookmark not defined.
Bookmark not defined.	
2.1.7 Algoritma Naive Bayes	Error! Bookmark not defined.
2.1.8 <i>Confusion Matrix</i> dan Metrik Evaluasi.....	Error! Bookmark not defined.
2.2 Teknologi yang digunakan	Error! Bookmark not defined.
2.2.1 Python	Error! Bookmark not defined.
2.2.2 Streamlit.....	Error! Bookmark not defined.
2.2.3 <i>Library</i> Python.....	Error! Bookmark not defined.
2.3 Metodologi Pengembangan Perangkat Lunak	Error! Bookmark not defined.
2.4 Kakas Pemodelan	Error! Bookmark not defined.
2.4.1 <i>Flowchart</i>	Error! Bookmark not defined.
2.4.2 <i>Data Flow Diagram</i> (DFD).....	Error! Bookmark not defined.
2.5 Metode Pengumpulan Data	Error! Bookmark not defined.
2.6 Pengolahan Data.....	Error! Bookmark not defined.
2.7 Penelitian Terkait.....	Error! Bookmark not defined.
 BAB III ANALISIS	 Error! Bookmark not defined.
3.1 Identifikasi Masalah dan Kesempatan	Error! Bookmark not defined.

3.2 Analisis Target Pengguna	Error! Bookmark not defined.
3.3 Daftar Spesifikasi Persyaratan	Error! Bookmark not defined.
3.3.1 Persyaratan Fungsional	Error! Bookmark not defined.
3.3.2 Persyaratan Non-Fungsional.....	Error! Bookmark not defined.
3.4 Analisis Penerapan Algoritma	Error! Bookmark not defined.
3.4.1 Tahap <i>Pre-Processing</i>	Error! Bookmark not defined.
3.4.2 Tahap <i>Processing</i>	Error! Bookmark not defined.
3.4.3 Tahap Validasi	Error! Bookmark not defined.
 BAB IV PERANCANGAN	Error! Bookmark not defined.
4.1 Perancangan Aplikasi	Error! Bookmark not defined.
4.2 Perancangan Diagram	Error! Bookmark not defined.
4.3 Perancangan Antarmuka.....	Error! Bookmark not defined.
 BAB V IMPLEMENTASI	Error! Bookmark not defined.
5.1 Lingkungan Implementasi.....	Error! Bookmark not defined.
5.1.1 Perangkat Keras (<i>Hardware</i>)	Error! Bookmark not defined.
5.1.2 Perangkat Lunak (<i>Software</i>)	Error! Bookmark not defined.
5.2 Batasan Implementasi	Error! Bookmark not defined.
5.3 Implementasi Antarmuka Aplikasi	Error! Bookmark not defined.
5.4 Implementasi Modul Program Utama.....	Error! Bookmark not defined.
5.4.1 Modul Autentikasi dan Keamanan Kredensial.....	Error! Bookmark not defined.
5.4.2 Modul Pra-Pemrosesan Bahasa Alami (NLP).....	Error! Bookmark not defined.
5.4.3 Modul Pelatihan Algoritma Naive Bayes.....	Error! Bookmark not defined.
5.4.4 Modul Inferensi dan Prediksi Sentimen	Error! Bookmark not defined.
 BAB VI PENGUJIAN	Error! Bookmark not defined.
6.1 Tujuan Pengujian.....	Error! Bookmark not defined.
6.2 Kriteria Pengujian	Error! Bookmark not defined.
6.3 Pengujian Fungsional (<i>Black Box Testing</i>).....	Error! Bookmark not defined.
6.4 Pengujian Algoritma Klasifikasi (Performa Model).....	Error! Bookmark not defined.
6.4.1 Pengujian Rasio Data Latih 70% dan Data Uji 30% (70:30)	Error! Bookmark not defined.
6.4.2 Pengujian Rasio Data Latih 80% dan Data Uji 20% (80:20)	Error! Bookmark not defined.
6.4.3 Pengujian Rasio Data Latih 90% dan Data Uji 10% (90:10)	Error! Bookmark not defined.
6.5 Analisis Hasil Pengujian	Error! Bookmark not defined.
 BAB VII KESIMPULAN DAN SARAN	Error! Bookmark not defined.
7.1 Kesimpulan	Error! Bookmark not defined.
7.2 Saran.....	Error! Bookmark not defined.
 DAFTAR PUSTAKA	Error! Bookmark not defined.

UKDLSM

DAFTAR TABEL

Tabel 2.1 <i>Library</i> Python yang Digunakan	Error! Bookmark not defined.
Tabel 2.2 Simbol <i>Flowchart</i> [19]	Error! Bookmark not defined.
Tabel 2.3 Simbol DFD Yourdan dan De Marco [19]	Error! Bookmark not defined.
Tabel 2.4 Distribusi <i>Dataset</i> Skenario 1 (Data Tidak Seimbang)	Error! Bookmark not defined.
Tabel 2.5 Penelitian Terkait	Error! Bookmark not defined.
Tabel 3.1 Persyaratan Fungsional Aplikasi	Error! Bookmark not defined.
Tabel 3.2 Distribusi <i>Dataset</i> Skenario 1 (Data Tidak Seimbang)	Error! Bookmark not defined.
Tabel 3.3 Distribusi <i>Dataset</i> Skenario 2 (Data Seimbang)	Error! Bookmark not defined.
Tabel 3.4 Contoh Hasil Pra-Pemrosesan Teks	Error! Bookmark not defined.
Tabel 3.5 Data Latih untuk Demonstrasi Perhitungan Algoritma	Error! Bookmark not defined.
Tabel 3.6 Frekuensi Kemunculan Kata per Kelas Sentimen dari Data Latih.	Error! Bookmark not defined.
Tabel 3.7 Nilai Probabilitas Kondisional $P(w_i C_k)$ dengan Laplace Smoothing ($\alpha = 1,0$)	Error! Bookmark not defined.
Tabel 3.8 Rekap Nilai Log Posterior untuk Data Uji "bagus crash"	Error! Bookmark not defined.
Tabel 5.1 Spesifikasi Perangkat Keras	Error! Bookmark not defined.
Tabel 5.2 Daftar Perangkat Lunak Yang Digunakan	Error! Bookmark not defined.
Tabel 6.1 <i>Black Box Testing</i>	Error! Bookmark not defined.
Tabel 6.2 Hasil Pengujian Model dengan Rasio 70:30 (Tanpa <i>Random State</i>)	Error! Bookmark not defined.
Tabel 6.3 Hasil Pengujian Model dengan Rasio 80:20 (Tanpa <i>Random State</i>)	Error! Bookmark not defined.
Tabel 6.4 Hasil Pengujian Model dengan Rasio 90:10 (Tanpa <i>Random State</i>)	Error! Bookmark not defined.

DAFTAR GAMBAR

Gambar 2.1	Metode <i>Waterfall</i> [13]	Error! Bookmark not defined.
Gambar 4.1	Arsitektur Aplikasi	Error! Bookmark not defined.
Gambar 4.2	<i>Flowchart</i> Pra-Pemrosesan Teks	Error! Bookmark not defined.
Gambar 4.3	<i>Flowchart</i> Prediksi Sentimen	Error! Bookmark not defined.
Gambar 4.4	DFD Level 0	Error! Bookmark not defined.
Gambar 4.5	DFD Level 1	Error! Bookmark not defined.
Gambar 4.6	Halaman <i>Login User</i>	Error! Bookmark not defined.
Gambar 4.7	Halaman Beranda	Error! Bookmark not defined.
Gambar 4.8	Halaman Unggah <i>Dataset</i>	Error! Bookmark not defined.
Gambar 4.9	Halaman Pra-Pemrosesan	Error! Bookmark not defined.
Gambar 4.10	Halaman Pelatihan Model	Error! Bookmark not defined.
Gambar 4.11	Halaman Evaluasi Model	Error! Bookmark not defined.
Gambar 4.12	Halaman Prediksi Sentimen	Error! Bookmark not defined.
Gambar 4.13	Halaman Riwayat Prediksi	Error! Bookmark not defined.
Gambar 4.14	Halaman Kelola Pengguna	Error! Bookmark not defined.
Gambar 5.1	Halaman Autentikasi (<i>Login</i>)	Error! Bookmark not defined.
Gambar 5.2	Halaman Beranda	Error! Bookmark not defined.
Gambar 5.3	Halaman Manajemen <i>Dataset</i>	Error! Bookmark not defined.
Gambar 5.4	Halaman Pra-Pemrosesan Teks	Error! Bookmark not defined.
Gambar 5.5	Halaman Pelatihan Model	Error! Bookmark not defined.
Gambar 5.6	Halaman Evaluasi Model	Error! Bookmark not defined.
Gambar 5.7	Halaman Prediksi Sentimen	Error! Bookmark not defined.
Gambar 5.8	Tampilan Riwayat Prediksi	Error! Bookmark not defined.
Gambar 5.9	Halaman Kelola Pengguna	Error! Bookmark not defined.
Gambar 5.10	Modul Autentikasi dan Keamanan Kredensial	Error! Bookmark not defined.
Gambar 5.11	Modul Pra-Pemrosesan Bahasa Alami (NLP)	Error! Bookmark not defined.
Gambar 5.12	Modul Pelatihan Algoritma Naive Bayes	Error! Bookmark not defined.
Gambar 5.13	Modul Inferensi dan Prediksi Sentimen	Error! Bookmark not defined.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi informasi yang terjadi secara masif dalam beberapa tahun terakhir telah mendorong perubahan besar dalam pola interaksi sosial masyarakat. Berbagai platform digital bermunculan sebagai wadah baru bagi pengguna untuk berbagi informasi, mengekspresikan pendapat, serta terhubung satu sama lain secara daring. Salah satu platform yang belakangan ini menarik perhatian adalah Lemon8, sebuah aplikasi media sosial berbasis konten gaya hidup yang dikembangkan oleh ByteDance. Aplikasi ini pertama kali diluncurkan di Jepang pada tahun 2020 dan kemudian berkembang pesat ke berbagai negara, termasuk Indonesia, dengan fokus utama pada konten gaya hidup, seperti *fashion*, kecantikan, dan kuliner. Sejak dirilis, Lemon8 tercatat telah mencapai sekitar 5 juta pengguna aktif setiap bulan di seluruh dunia, termasuk Indonesia, dengan total unduhan mencapai kurang lebih 15 juta kali [1]. Angka pengguna yang dipilah khusus untuk Indonesia memang belum dipublikasikan secara terbuka oleh pihak pengembang, tetapi tingkat popularitas Lemon8 di Tanah Air tergambar dari posisinya di App Store. Per Juni 2023, aplikasi ini sudah menduduki peringkat ketiga aplikasi gaya hidup gratis terpopuler di Indonesia [1].

Tingginya minat masyarakat terhadap Lemon8 turut berdampak pada meningkatnya jumlah ulasan pengguna yang diunggah di platform distribusi aplikasi seperti Google Play Store. Ulasan-ulasan tersebut mengandung berbagai bentuk ekspresi, mulai dari kepuasan, keluhan, saran, hingga perbandingan dengan aplikasi lain. Namun, volume data yang besar dan sifatnya yang tidak terstruktur membuat proses memahami isi keseluruhan ulasan secara konvensional menjadi tidak efisien dan membutuhkan waktu yang sangat lama. Oleh karena itu, diperlukan sebuah pendekatan komputasi menggunakan teknik penambangan teks (*text mining*) dan Analisis Sentimen untuk mengotomatisasi pengelompokan opini publik secara instan.

Analisis sentimen merupakan salah satu teknik dalam bidang pengolahan bahasa alami yang berfungsi untuk mengidentifikasi dan mengklasifikasikan opini

yang terkandung dalam sebuah teks ke dalam kategori tertentu, seperti positif, negatif, atau netral. Dalam konteks ulasan aplikasi, penerapan analisis sentimen dinilai mampu memberikan gambaran menyeluruh mengenai tingkat kepuasan pengguna secara efisien, sehingga hasil analisisnya dapat dijadikan dasar pengambilan keputusan oleh pihak pengembang dalam meningkatkan kualitas produk [2].

Di antara berbagai algoritma yang umum digunakan dalam analisis sentimen, Naive Bayes menonjol sebagai salah satu pilihan yang paling banyak diterapkan dalam penelitian berbasis teks. Algoritma ini bekerja berdasarkan teori probabilitas dengan asumsi bahwa setiap kata dalam dokumen bersifat independen satu sama lain. Meskipun asumsi tersebut tergolong sederhana, Naive Bayes terbukti mampu menghasilkan tingkat akurasi yang kompetitif dalam tugas klasifikasi teks. Selain itu, efisiensinya dalam memproses data berukuran besar dengan beban komputasi yang relatif rendah menjadikannya lebih unggul dibandingkan beberapa algoritma lain [3]. Penerapan Naive Bayes pada data ulasan aplikasi telah terbukti menghasilkan performa yang baik, dengan rata-rata akurasi mencapai 90%, presisi 98%, dan *recall* 81% [4]. Efisiensinya dalam memproses data berukuran besar dengan beban komputasi yang relatif rendah menjadikan algoritma ini lebih unggul dibandingkan beberapa pendekatan lain dalam skenario klasifikasi teks [5].

Sejumlah penelitian terdahulu telah membuktikan keberhasilan kombinasi antara analisis sentimen dan algoritma Naive Bayes dalam mengolah ulasan pengguna dari Google Play Store pada berbagai aplikasi populer, seperti aplikasi perbankan [6], platform rekrutmen kerja [7], hingga aplikasi pemilu [4]. Akan tetapi, hingga saat ini belum ditemukan penelitian yang secara spesifik mengarahkan kajiannya pada aplikasi Lemon8. Celah penelitian inilah yang menjadi landasan dilakukannya studi ini, mengingat pemahaman terhadap sentimen pengguna Lemon8 memiliki relevansi yang tinggi di tengah pesatnya perkembangan platform tersebut di Indonesia.

Berdasarkan uraian di atas, penelitian ini mengambil judul "Analisis Sentimen Ulasan Aplikasi Lemon8 di Google Play Store Menggunakan Algoritma Naive Bayes". Melalui penelitian ini, diharapkan dapat dihasilkan sebuah model

klasifikasi sentimen yang akurat sekaligus memberikan wawasan yang berguna mengenai persepsi nyata pengguna terhadap aplikasi Lemon8.

1.2 Rumusan Masalah

Bagaimana membangun sebuah aplikasi analisis sentimen ulasan aplikasi Lemon8 di Google Play Store menggunakan algoritma Naive Bayes?

1.3 Tujuan Tugas Akhir

Membangun sebuah aplikasi analisis sentimen ulasan aplikasi Lemon8 di Google Play Store menggunakan algoritma Naive Bayes.

1.4 Manfaat Tugas Akhir

1. Bagi *User* (Orang awam)

Membantu calon pengguna dalam menentukan apakah aplikasi Lemon8 layak diunduh berdasarkan rangkuman sentimen (positif/negatif) dari ribuan ulasan yang sudah diproses secara otomatis, bukan sekadar melihat rating bintang.

2. Bagi Peneliti

Menguji tingkat akurasi algoritma Naive Bayes dalam menangani data ulasan bahasa Indonesia yang sering kali menggunakan bahasa tidak baku pada Google Play Store.

1.5 Batasan Masalah

Berikut beberapa batasan masalah pada penelitian ini.

1. Data ulasan yang digunakan dalam penelitian ini hanya bersumber dari Google Play Store, dan dibatasi pada ulasan berbahasa Indonesia.
2. Jumlah *dataset* hanya 3914 data, yang terbagi dari 2917 positif, 620 negatif dan 377 netral.
3. Klasifikasi hanya dibagi ke dalam tiga label, dengan kriteria: Positif (pujian, rasa puas, rekomendasi aplikasi), Negatif (keluhan eror, rasa frustrasi, ketidakpuasan), dan Netral (pertanyaan umum, saran fitur tanpa emosi, atau informasi biasa).

4. Data ulasan diasumsikan murni ditulis oleh manusia. Aplikasi tidak menyediakan fitur pendeteksian akun otomatis (*bot/spam*) sehingga seluruh data yang berhasil ditarik dari Google Play Store dianggap sebagai opini yang benar.
5. Aplikasi analisis sentimen ini berfokus pada ekstraksi fitur berupa teks kata murni (karakter alfabet). Karakter khusus berupa angka, tanda baca, tautan (URL), dan karakter visual seperti *em*
6. *Emoticon* atau emoji yang terdapat di dalam ulasan pengguna tidak dievaluasi. Seluruh karakter *non*-teks tersebut akan dihapus secara otomatis pada tahapan pra-pemrosesan data (tahap *cleaning*) dan tidak dihitung ke dalam pembobotan fitur (TF-IDF) untuk memprediksi kelas sentimen.

1.6 Metodologi Penelitian

Metode yang digunakan pada penelitian ini meliputi beberapa tahapan sebagai berikut.

1. Identifikasi Masalah
Tahap awal ini akan disusun latar belakang dengan mengidentifikasi masalah berdasarkan topik penelitian yang akan menghasilkan identifikasi masalah dan kebutuhan.
2. Penentuan Tujuan
Tahap kedua ini akan menentukan tujuan dari penelitian ini dan proses pengumpulan data yang diperlukan.
3. Studi Literatur
Tahap ketiga ini penulis mengumpulkan teori-teori pendukung yang berkaitan dengan penelitian ini beserta pembahasan dan metodologi penelitian.
4. Pengumpulan Data dan Analisis
Tahap keempat ini adalah mengidentifikasi masalah, menentukan target pengguna dan menganalisis penerapan algoritma Naive Bayes.

5. Perancangan

Tahap kelima ini berisi perancangan aplikasi, antara lain rancangan alur aplikasi melalui beberapa diagram pemodelan, rancangan basis data, rancangan tampilan antarmuka aplikasi dan perancangan modul program.

6. Implementasi

Tahap ini akan mengimplementasikan hasil rancangan sebelumnya.

7. Pengujian

Pada tahapan ini akan dilakukan pengujian terhadap aplikasi untuk memastikan aplikasi sudah sesuai spesifikasi dan berjalan dengan baik.

1.7 Sistematika Penulisan

Laporan Tugas Akhir ini terdiri dari beberapa bagian sebagai berikut.

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang, rumusan masalah, tujuan dan manfaat dari Tugas Akhir, batasan masalah, metode penelitian dan sistematika penulisan.

BAB II STUDI PUSTAKA

Bab ini menjelaskan tentang teori pendukung terkait topik penelitian Tugas Akhir, teknologi yang digunakan, metodologi pengembangan perangkat lunak, kakas pemodelan dan penelitian terkait.

BAB III ANALISIS

Bab ini berisi penjelasan tentang identifikasi masalah, spesifikasi persyaratan dan analisis penerapan algoritma Naive Bayes

BAB IV PERANCANGAN

Bab ini berisi rancangan basis data, tampilan antarmuka aplikasi dan rancangan modul program.

BAB V IMPLEMENTASI

Bab ini mengimplementasikan keseluruhan dari hasil rancangan.

BAB VI PENGUJIAN

Bab ini merupakan pengujian dari aplikasi untuk memastikan aplikasi yang dirancang sudah sesuai dengan spesifikasi persyaratan dan berjalan dengan baik.

BAB VII KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan laporan dan aplikasi yang telah dibangun dan saran terhadap pengembangan aplikasi selanjutnya.

UKDLSM