

**ANALISIS SENTIMEN ASPECT-BASED ULASAN LINE
MESSENGER PADA GOOGLE PLAY STORE
MENGGUNAKAN ALGORITMA INDOBERT DAN LDA**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



disusun oleh

ATQIYA TRIANDA PUTRA ANUGRAH

22.11.5083

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

DAFTAR ISI

Halaman Judul	1
DAFTAR ISI	2
BAB I PENDAHULUAN.....	4
1.1 Latar Belakang	4
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
1.5.1 Manfaat Teoritis.....	6
1.5.2 Manfaat Praktis.....	6
1.6 Sistematika Penulisan.....	7
BAB II TINJAUAN PUSTAKA	8
2.1 Studi Literatur	8
2.2 Dasar Teori.....	10
2.2.1 Analisis Sentimen.....	10
2.2.2 Analisis Sentimen Berbasis Aspek (ABSA).....	11
2.2.3 BERT	11
2.2.4 IndoBERT.....	11
2.2.5 Latent Dirichlet Allocation (LDA)	11
2.2.6 Data Cleaning	11
2.2.7 Case Folding.....	12
2.2.8 Normalisasi.....	12
2.2.9 Tokenization	12
2.2.10 Stopword Removal.....	12
2.2.11 Stemming	12
2.2.12 TF-IDF (Term Frequency-Inverse Document Frequency)	13
Bab III METODE PENELITIAN	14
3.1 Objek Penelitian.....	14
3.2 Alur Penelitian	14
3.3 Alat dan Bahan.....	16
3.4 Teknik Preprocessing	16

3.5	Teknik Evaluasi.....	16
	Bab IV HASIL DAN PEMBAHASAN	17
4.1	Data Scraping.....	17
4.2	Preprocessing	17
4.2.1	Data Cleaning.....	17
4.2.2	Case Folding.....	18
4.2.2	Normalization.....	18
4.2.3	Stopword	19
4.2.4	Stemming	19
4.2.5	Tokenization	20
4.2.5	TF-IDF	20
4.3	Ekstraksi Topik Dan Evaluasi Model LDA	21
4.5	Modelling Indobert	25
4.5.1	Komponen Utama Arsitektur Model	25
4.5.2	Spesifikasi Teknis Model	26
4.5.3	Multi-Topic Classification.....	26
4.5.4	Tokenization dan Preprocessing.....	26
4.5.5	Fine-Tuning Strategy.....	27
4.5.6	Evaluasi Model IndoBERT	27
4.6	Deployment & Uji Model Dengan Data Nyata Baru	28
4.7	Pembahasan	30
	BAB V PENUTUP	32
5.1	Kesimpulan.....	32
5.2	Saran	33
	DAFTAR PUSTAKA	34

BAB I

PENDAHULUAN

1.1 Latar Belakang

LINE Messenger adalah salah satu aplikasi *Messaging* dan *Social Media* yang pernah popular di Indonesia. Menurut laporan “Indotelko.com”[1], pada bulan September tahun 2023, LINE memiliki 196 juta pengguna aktif secara global dengan jangka waktu bulanan. Indonesia sendiri memiliki kontribusi total jumlah pengguna sebesar 179 juta pengguna, hal ini menjadikan Indonesia sebagai salah satu pasar utama LINE di dunia. Tidak hanya itu, “Liputan 6” juga menyebutkan bahwa sepanjang tahun 2024, pengguna LINE di Indonesia mengirimkan lebih dari 500 juta stiker, Fakta ini menunjukkan bahwa LINE tetap mempertahankan jumlah pengguna yang besar dan aktif di Indonesia [2]. Popularitas ini memperkuat posisi LINE sebagai salah satu aplikasi *Messaging* dan *Social Media* yang pernah populer di Indonesia.

Namun, dalam beberapa tahun terakhir, LINE mulai mengurangi fitur-fiturnya di Indonesia. Sejak 2022, beberapa fitur utama telah dihentikan, seperti layanan berita Line Today pada 6 Juli 2022 dan fitur obrolan publik OpenChat pada 20 Juli 2022, diikuti dengan penutupan resmi layanan media sosial konten pendek VOOM per 14 Februari 2025[3]. Penutupan ini disebut sebagai bagian dari strategi bisnis baru LINE dikarenakan fitur-fitur tersebut kalah bersaing di pasar lokal. Penutupan fitur-fitur ini menimbulkan beragam pendapat pengguna terhadap aplikasi, ada beberapa yang mungkin merasa kehilangan fitur yang sebelumnya mereka nikmati, namun ada juga yang mungkin merasa bahwa penutupan fitur ini merupakan perubahan yang bagus bagi aplikasi.

Untuk memahami pendapat pengguna terhadap LINE Messenger di tengah perubahan ini, analisis sentimen berbasis aspek menjadi sangat penting, dengan data ulasan pengguna dari Google Play Store sebagai sumber yang berharga untuk menganalisis opini-opini tentang fitur-fitur dan performa aplikasi. Mengingat volume ulasan yang besar, pendekatan machine learning (ML) diperlukan untuk efisiensi dan akurasi. Menurut IBM, analisis sentimen modern menggabungkan pemrosesan bahasa alami (NLP) dan ML, yang lebih adaptif dan mampu menangani skenario kompleks dibanding metode tradisional secara manual[4]. Dalam konteks bahasa Indonesia, model khusus seperti IndoBERT sangat dibutuhkan karena mampu memahami nuansa lokal sebagai varian BERT monolingual yang dilatih pada Bahasa Indonesia, sehingga klasifikasi sentimen lebih tepat[5]. Dibandingkan dengan algoritma lain yang bersifat non-transformer seperti SVM, KNN dan

Random Forest Algoritma transformer seperti BERT memiliki performa yang lebih bagus secara metrik evaluasi [6][7]. Multilingual BERT juga menjadi salah satu pilihan selain IndoBERT, namun IndoBERT memiliki performa yang lebih bagus jika dinilai secara metrik saat dibandingkan dengan mBERT [8]. Sementara itu, Latent Dirichlet Allocation (LDA) digunakan untuk mengekstraksi aspek pada ulasan secara otomatis, mengidentifikasi pola hubungan kata dan aspek-aspek yang menjadi perhatian [4]. Kombinasi IndoBERT dan LDA diharapkan memberikan hasil yang dapat dipahami dengan mudah tanpa perlu menggunakan metode tradisional [4][5]. Penelitian ini bertujuan memberikan gambaran jelas tentang bagaimana pengguna Indonesia menilai LINE, yang dapat membantu strategi LINE atau aplikasi *Messaging* lain di masa depan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Apa saja aspek-aspek Line Messenger yang diidentifikasi melalui analisis berbasis machine learning dari ulasan pengguna di Google Play Store?
2. Bagaimana sentimen pengguna terhadap layanan aplikasi Line Messenger yang dihasilkan dari model analisis sentimen berbasis machine learning, berdasarkan ulasan di Google Play Store?
3. Bagaimana sentimen pengguna terhadap masing-masing aspek Line Messenger yang teridentifikasi melalui teknik analisis berbasis machine learning, berdasarkan ulasan di Google Play Store?
4. Bagaimana Performa IndoBERT sebagai model NLP pada kasus Aspect Based Sentiment Analysis dengan menilai berdasarkan Metrik Evaluasi Coherence, F1 Score, Precision Dan Recall

1.3 Batasan Masalah

Untuk memfokuskan penelitian ini, beberapa batasan masalah ditetapkan sebagai berikut:

1. Data yang digunakan adalah ulasan pengguna Line Messenger dalam bahasa Indonesia di Google Play Store pada rentang waktu 2022-2025 sebanyak 7.000 ulasan untuk Modelling LDA dan 950 untuk IndoBERT.
2. Analisis sentimen dibatasi pada tiga kategori yaitu positif, negatif, dan netral.

3. Penelitian tidak membahas faktor eksternal (seperti kebijakan pemerintah atau strategi kompetitor) yang mungkin memengaruhi alasan sentimen negatif.
4. Algoritma yang digunakan adalah IndoBERT untuk klasifikasi sentimen dan LDA untuk ekstraksi aspek.
5. Aspek yang diteliti dibatasi pada kategori yang muncul secara dominan dari hasil pemodelan LDA.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Mengidentifikasi aspek-aspek Line Messenger yang dikomentari oleh pengguna dalam ulasan Google Play Store pada Line Messenger menggunakan teknik ekstraksi aspek berbasis machine learning.
2. Menganalisis sentimen pengguna terhadap masing-masing aspek Line Messenger dengan menerapkan model analisis sentimen berbasis machine learning pada ulasan di Google Play Store.
3. Menganalisis sentimen keseluruhan pengguna terhadap layanan aplikasi Line Messenger dengan memanfaatkan metode analisis sentimen berbasis aspek menggunakan machine learning, pada ulasan di Google Play Store.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1.5.1 Manfaat Teoritis

1. Memberikan kontribusi pada pengembangan metode analisis sentimen berbasis aspek untuk ulasan aplikasi dalam bahasa Indonesia.
2. Memperkaya wawasan tentang efektivitas kombinasi algoritma IndoBERT dan LDA dalam analisis sentimen berbasis aspek.

1.5.2 Manfaat Praktis

1. Bagi tim pengembang Line Messenger, hasil penelitian ini dapat menjadi masukan berharga untuk meningkatkan kualitas aplikasi dan memperbaiki aspek-aspek yang mendapat sentimen negatif dari pengguna.
2. Bagi pelaku industri aplikasi *Messaging*, penelitian ini dapat menjadi referensi dalam memahami preferensi dan kebutuhan pengguna Indonesia.

3. Bagi pengguna aplikasi, penelitian ini dapat memberikan informasi lengkap tentang kelebihan dan kekurangan Line Messenger dibandingkan dengan aplikasi pesan instan lainnya.
4. Bagi peneliti lain, penelitian ini dapat dijadikan referensi untuk melakukan analisis serupa pada aplikasi lain di Google Play Store.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN, Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA, Bab ini berisi tinjauan literatur dari penelitian terdahulu yang relevan serta dasar teori yang mendukung penelitian. Tinjauan pustaka bertujuan untuk memberikan landasan konseptual dan teoritis.

BAB III METODE PENELITIAN, Bab ini berisi tentang penjelasan metode yang digunakan dalam penelitian, termasuk jenis dan sumber data, teknik pengumpulan data, serta tahapan-tahapan penelitian mulai dari preprocessing data, ekstraksi aspek dengan LDA, hingga evaluasi model IndoBERT.

BAB IV HASIL DAN PEMBAHASAN, Bab ini berisi hasil dari analisis yang dilakukan, baik dari hasil topic modeling maupun klasifikasi sentimen. Selain itu, bab ini juga membahas evaluasi kinerja model berdasarkan metrik seperti coherence, presisi, dan recall, serta interpretasi hasil terhadap aspek-aspek penting dari Line Messenger.

BAB V PENUTUP, Bab ini berisi kesimpulan dari hasil penelitian serta saran-saran yang dapat dijadikan acuan untuk pengembangan penelitian selanjutnya di masa mendatang.

BAB II

TINJAUAN PUSTAKA

2.1 Studi Literatur

Awal studi ini didasari oleh penelitian “Analisis sentimen terhadap kualitas pelayanan aplikasi GoJek menggunakan metode Naive Bayes Classifier” [6]. Dengan mengolah 800 ulasan di Google Play Store (640 untuk pelatihan, 160 untuk pengujian), peneliti melakukan serangkaian pra-pemrosesan case folding, tokenizing, filtering, dan stemming peneliti kemudian mengekstrak fitur TF-IDF dan menampilkan hasilnya lewat wordcloud. Meski berhasil mencapai akurasi 68 %, metode Naive Bayes masih kesulitan menangkap nuansa Bahasa Indonesia.

Di sisi lain, Putri et al. dalam “Analisis Sentimen dan Pemodelan Ulasan Aplikasi AdaKami Menggunakan Algoritma SVM dan KNN” [7] mengeksplorasi metode klasik dengan kerangka SEMMA pada 1.000 ulasan manual. SVM terbukti paling unggul—akurasi, presisi, recall, dan F1-Score masing-masing 93 % pada pembagian data 90:10—Penelitian ini yang menjadi dasar pemilihan SVM sebagai salah satu algoritma yang akan dibandingkan di penelitian Ini.

Lalu memasuki ranah mBERT dan IndoBERT, penelitian “Analisis Sentimen Pengguna Platform Media Sosial X pada Topik Pemilihan Presiden 2024 Menggunakan Perbandingan Model Monolingual dan Multilingual BERT” [8] memperlihatkan keunggulan IndoBERT dibanding mBERT. Dari 10.140 tweet berbahasa Indonesia, IndoBERT mencatat akurasi 84 % dan F1-Score 78 %.

Penelitian “Performance analysis of IndoBERT for sentiment classification in Indonesian hotel review data” [9] mengangkat tantangan ketidakseimbangan data. Meskipun akurasi globalnya tinggi (92,52 %), model ini gagal mendeteksi dengan baik ulasan netral dan negatif, sehingga penulis merekomendasikan penggunaan SMOTE atau penyesuaian bobot kelas untuk mengurangi bias.

Terakhir, “Aspect-Based Sentiment Analysis with LDA and IndoBERT Algorithm on Mental Health App: Riliv” [10] menunjukkan kombinasi LDA dan IndoBERT. Dari 3.068 ulasan, pendekatan ini meraih akurasi 95 % dan mampu mengatasi ketidakseimbangan data lewat SMOTE. Tidak hanya itu penulis juga mencoba menggunakan klasifikasi non transformer yang salah satunya adalah Random Forest dan meraih akurasi tertinggi dibawah indoBERT dengan akurasi

94%, Hasil ini memberi pondasi yang kuat untuk menerapkan IndoBERT dan Random Forest sebagai algoritma yang akan dibandingkan.

Judul Penelitian	Nama Penulis	Tahun	Hasil Penelitian	Perbandingan dengan Penelitian Ini
“Analisis sentimen terhadap kualitas pelayanan aplikasi GoJek menggunakan metode Naive Bayes Classifier”	Indarwati dan Februariyanti	2023	Akurasi 68% (Naive Bayes+TF-IDF);	Metode tradisional (Naive Bayes) tanpa ABSA; akurasi lebih rendah dibandingkan pendekatan IndoBERT+LDA dalam penelitian ini.
“Analisis Sentimen dan Pemodelan Ulasan Aplikasi AdaKami menggunakan algoritma SVM dan KNN”	Putri et al.	2023	SVM akurasi ~93% (lebih tinggi dari KNN); klasifikasi umum,	Pendekatan tradisional (SVM) tanpa ABSA; akurasi tinggi namun berbeda dengan penelitian ini yang menggunakan ABSA (LDA) untuk analisis aspek.
“Analisis sentimen pengguna platform media sosial X pada topik pemilihan presiden	Nurhasiyah et al.	2024	IndoBERT unggul dengan F1-score 78%; fokus politik,	Penggunaan IndoBERT sama dengan penelitian ini, namun hanya pada konteks politik tanpa ABSA;

2024 menggunakan perbandingan model monolingual dan multilingual BERT”				
“Performance analysis of IndoBERT for sentiment classification in Indonesian hotel review data”	Singgalen	2025	Akurasi global 92%; rendah pada kelas minoritas (data imbalanced).	Menggunakan IndoBERT seperti penelitian ini, tetapi tanpa ABSA
“Aspect-Based Sentiment Analysis with LDA and IndoBERT Algorithm on Mental Health App: Riliv”	Aryanti et al.	2025	Akurasi 95% dengan IndoBERT + LDA; berhasil menangani ketidakseimbangan data.	Sangat relevan: menggunakan IndoBERT+LDA untuk ABSA di Riliv; akurasi tinggi (95%) mendukung pendekatan penelitian ini.

2.2 Dasar Teori

2.2.1 Analisis Sentimen

Analisis sentimen adalah teknik dalam pengolahan bahasa alami yang mengevaluasi dan mengidentifikasi nada emosional atau sikap yang tersirat dalam data teks [11]. Dalam proses ini, kalimat atau dokumen dipindai untuk menentukan apakah teks tersebut mengandung sentimen positif, negatif, atau netral. Metode ini penting karena mampu mengekstrak insight yang berharga dari kumpulan data teks yang besar, membantu penganalisa memahami pendapat pelanggan untuk pengambilan keputusan yang lebih baik [11].

2.2.2 Analisis Sentimen Berbasis Aspek (ABSA)

Analisis sentimen berbasis aspek (ABSA) adalah pendekatan analisis sentimen yang lebih rinci (fine-grained) yang dirancang untuk mengidentifikasi aspek-aspek spesifik beserta opini terkait yang terkandung dalam suatu teks[12]. Dalam ABSA, langkah utama meliputi ekstraksi aspek dan klasifikasi sentimen untuk setiap aspek tersebut[13]. Dengan demikian, ABSA memungkinkan pemisahan sentimen pada tingkat aspek tertentu, memberikan wawasan yang lebih terarah dibandingkan analisis sentimen konvensional yang hanya menilai keseluruhan kalimat atau dokumen[12][13].

2.2.3 BERT

BERT (Bidirectional Encoder Representations from Transformers) adalah model representasi bahasa berbasis arsitektur transformer yang unggul dalam berbagai tugas NLP. Model ini dilatih secara dua-arah (bidirectional) menggunakan teknik masked language modeling, sehingga mampu menangkap konteks kata dari kedua sisi kalimat [14]. Keunggulan BERT dalam 9 memahami konteks kontekstual dua-arah ini telah mendorong adopsi luasnya dalam penelitian dan aplikasi NLP modern.

2.2.4 IndoBERT

IndoBERT adalah versi pra-latih model BERT monolingual untuk Bahasa Indonesia. Model ini dikembangkan dengan melatih arsitektur BERT pada korpus teks Bahasa Indonesia yang sangat besar (Indo4B, 23,43 GB) [15]. Pelatihan dua-arah yang digunakan oleh IndoBERT memungkinkan embedding kata mengandung informasi semantik kontekstual dari kedua sisi teks, sehingga model ini dapat menangkap makna kata dalam konteks Bahasa Indonesia dengan lebih baik.

2.2.5 Latent Dirichlet Allocation (LDA)

Menurut Dami dan Alimardani [16], Latent Dirichlet Allocation (LDA) adalah teknik unsupervised untuk mengekstraksi informasi tematik dari sekumpulan dokumen. Inti dari LDA adalah mengasumsikan bahwa setiap dokumen dapat dianggap sebagai kombinasi acak dari beberapa topik laten, dan setiap topik didefinisikan sebagai distribusi probabilitas atas kata-kata[11]. Dengan demikian, LDA sering digunakan sebagai metode topic modeling untuk mengidentifikasi tema atau aspek utama yang muncul dalam data teks ulasan.

2.2.6 Data Cleaning

Data cleaning adalah proses menghilangkan "noise" atau data yang tidak relevan dari teks. Ini merupakan langkah penting dalam NLP untuk memastikan kualitas data yang akan diproses lebih lanjut [17].

2.2.7 Case Folding

Case folding adalah proses mengubah semua huruf dalam teks menjadi huruf kecil (lowercase). Ini dilakukan untuk menormalisasi teks sehingga kata yang sama dengan kapitalisasi berbeda (misalnya, "Line" dan "line") diperlakukan sebagai kata yang sama [18].

2.2.8 Normalisasi

Normalisasi teks melibatkan pengubahan angka, singkatan, atau slang menjadi bentuk standar. Tujuannya adalah untuk menyederhanakan variasi kata dalam ulasan sehingga kata-kata yang memiliki makna serupa dapat diidentifikasi secara konsisten [19].

2.2.9 Tokenization

Tokenization adalah proses memecah data kompleks seperti paragraf menjadi unit-unit sederhana yang disebut "token". Token dapat berupa kata, frasa, atau simbol. Ada dua jenis tokenization utama:

1. Sentence Tokenization: Memisahkan paragraf menjadi daftar kalimat.
2. Word Tokenization: Memisahkan kalimat menjadi daftar kata [18].

IndoBERT, sebagai model bahasa mutakhir berbasis BERT untuk Bahasa Indonesia, menggunakan tokenizer khusus yang secara otomatis menghasilkan sub-kata (subword) untuk model BERT. Ini berarti IndoBERT tidak hanya memecah teks menjadi kata-kata utuh, tetapi juga menjadi unit yang lebih kecil yang disebut subword. Pendekatan ini sangat efektif untuk menangani kata-kata yang tidak dikenal (Out-Of-Vocabulary/OOV) dan bahasa dengan morfologi yang kaya seperti Bahasa Indonesia [5].

2.2.10 Stopword Removal

Stopword adalah kata-kata umum dalam suatu bahasa (seperti "the", "a", "an", "in") yang membantu membentuk kalimat tetapi tidak memberikan makna signifikan dalam pemrosesan bahasa. Stopword removal adalah proses menghilangkan kata-kata ini dari teks untuk mengurangi dimensi data dan fokus pada kata-kata yang lebih informatif [17], [18].

2.2.11 Stemming

Stemming adalah teknik normalisasi di mana kata-kata yang terinfleksi (atau terkadang diturunkan) direduksi menjadi bentuk dasar atau akar kata. Proses ini bertujuan untuk mengurangi redundansi dengan mengubah daftar kata-kata yang di-token menjadi kata-kata akar yang lebih pendek. Contohnya, kata "memancing", "dipancing", dan "terpancing" akan direduksi menjadi "pancing" [18].

2.2.12 TF-IDF (Term Frequency-Inverse Document Frequency)

TF-IDF adalah statistik numerik yang bertujuan untuk mencerminkan seberapa penting sebuah kata bagi sebuah dokumen dalam koleksi atau korpus. TF-IDF umumnya lebih disukai daripada Bag of Words dalam model machine learning karena memberikan informasi tentang kata-kata yang lebih penting dan kurang penting [17].

1. **Term Frequency (TF):** Merupakan skor frekuensi kemunculan kata dalam dokumen saat ini [17].
2. **Inverse Document Frequency (IDF):** Merupakan skor seberapa jarang kata tersebut muncul di seluruh dokumen [17].

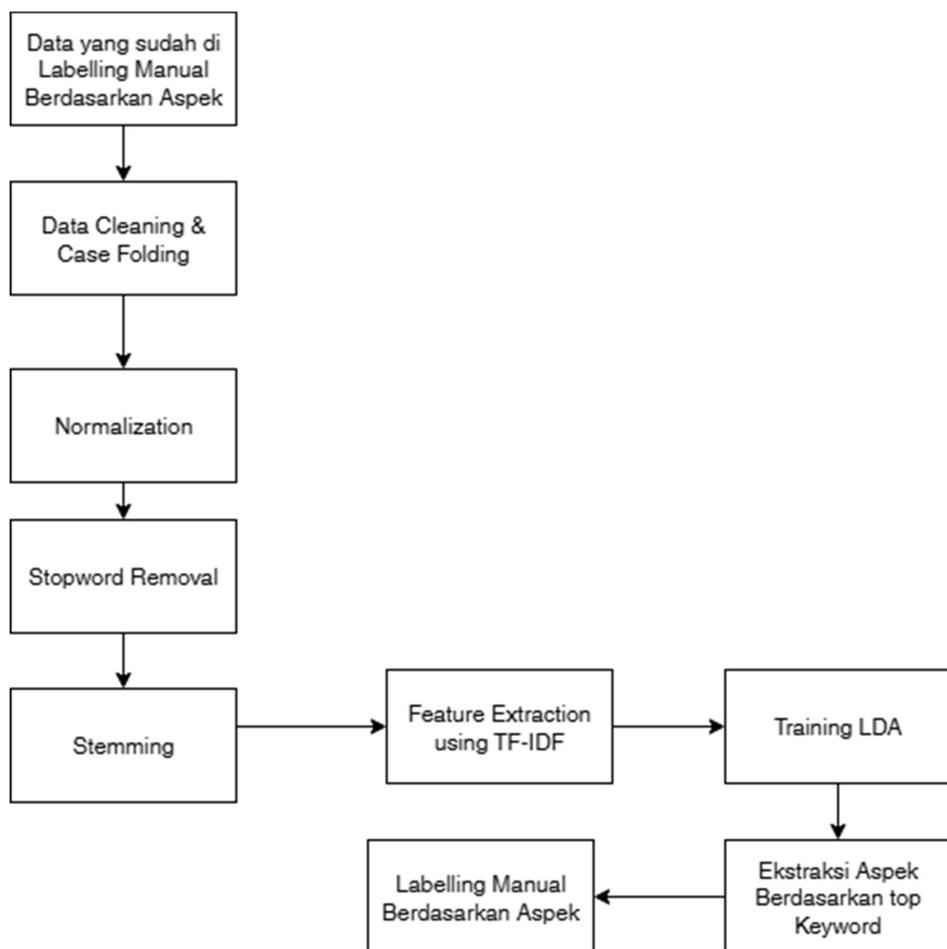
Bab III

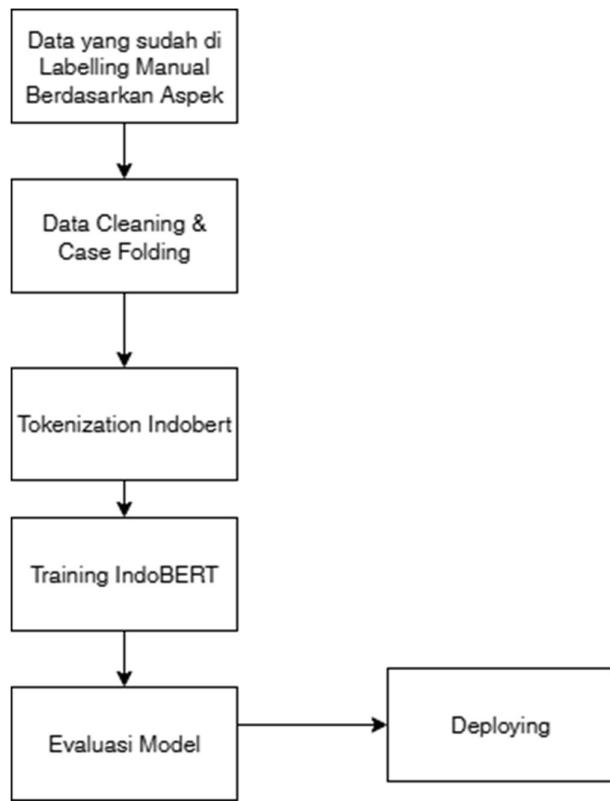
METODE PENELITIAN

3.1 Objek Penelitian

Objek penelitian adalah ulasan pengguna aplikasi Line Messenger yang terdapat di Google Play Store. Data ulasan dikumpulkan menggunakan library google-play-scraper pada Google Play Store, khususnya komentar berbahasa Indonesia. Sebanyak 7.000 data ulasan untuk memperoleh sampel yang representatif. Metode pengumpulan ini serupa dengan penelitian sebelumnya yang mengumpulkan 7.000 sampel ulasan pengguna dari Google Play Store via web scraping, Data kemudian dipotong sehingga 950 data untuk Training Indobert demi kemudahan komputasi.

3.2 Alur Penelitian





Gambar 1.1, 1.2 Alur Penelitian

Penelitian ini dimulai dengan pengumpulan data ulasan LINE Messenger dari Google Play Store menggunakan teknik *web scraping*. Data mentah yang diperoleh kemudian menjalani proses pembersihan (cleaning) dan *case folding* untuk menormalisasikan teks (misalnya mengubah semua huruf menjadi huruf kecil). Setelah pembersihan, dilakukan pengecekan apakah setiap ulasan sudah memiliki label per aspek. Jika data sudah berlabel per aspek, proses selanjutnya adalah tokenisasi dan pelatihan model IndoBERT.

Jika data belum berlabel aspek, tahap berikutnya adalah normalisasi teks, penghilangan *stopword*, dan *stemming* (misalnya dengan pustaka Sastrawi) untuk menyederhanakan variasi kata dalam ulasan. Dari teks yang sudah diproses ini, fitur diekstraksi menggunakan skema TF-IDF (Term Frequency–Inverse Document Frequency). Representasi TF-IDF ini selanjutnya digunakan untuk melatih model LDA. LDA (Latent Dirichlet Allocation) adalah teknik *topic modeling* yang berfungsi untuk menemukan topik-topik utama dalam sekumpulan dokumen. Model LDA menghasilkan distribusi topik berupa daftar kata kunci penting untuk setiap

topik. Berdasarkan kata kunci utama tersebut, peneliti melakukan pelabelan manual terhadap setiap ulasan sesuai aspek yang relevan.

Setelah seluruh data ulasan memiliki label aspek, data di-tokenisasi untuk masukan ke model IndoBERT. Model IndoBERT, yang merupakan model bahasa mutakhir berbasis BERT untuk Bahasa Indonesia, kemudian *di-fine tune* menggunakan data ulasan berlabel tersebut. Kinerja model yang telah dilatih dievaluasi menggunakan metrik akurasi, presisi, dan recall. Model terbaik kemudian diimplementasikan (*deployment*) untuk sistem analisis sentimen berbasis aspek pada aplikasi LINE Messenger.

3.3 Alat dan Bahan

Penelitian ini menggunakan perangkat lunak dan perangkat keras sebagai berikut:

- a. **Perangkat lunak:** Python, Sastrawi, gensim, scikit-learn, Huggingface Transformers, PyTorch, Weights & Biases (Wandb), pandas, numpy, Streamlit, Visual Studio Code, Google Colab, Microsoft Excel.
- b. **Perangkat keras:** CPU Intel Core i5-11400H, GPU NVIDIA GeForce RTX 3050, RAM 16 GB.

3.4 Teknik Preprocessing

Pra-pemrosesan teks terbagi berdasarkan jenis model. Untuk model LDA, dilakukan serangkaian tahap pra-pemrosesan umum pada teks Bahasa Indonesia, antara lain case folding (teks menjadi huruf kecil semua), pembersihan (cleaning), normalisasi (mengubah angka, singkatan, atau slang menjadi bentuk standar), penghapusan stopword (kata-kata umum yang tidak bermakna sentimen), dan stemming (mengambil akar kata menggunakan library Sastrawi). Tahapan ini sesuai dengan prosedur literatur yang banyak digunakan dalam NLP.

Sebaliknya, untuk model IndoBERT, hanya dilakukan tahapan case folding dan pembersihan sederhana, kemudian teks ditokenisasi menggunakan tokenizer IndoBERT yang secara otomatis menghasilkan sub-kata untuk model BERT.

3.5 Teknik Evaluasi

Metrik evaluasi yang digunakan adalah Coherence, presisi, recall, dan F1-score. Penggunaan Coherence dibanding accuracy untuk menilai performa klasifikasi sentimen karena dapat adanya class imbalance pada dataset

Bab IV

HASIL DAN PEMBAHASAN

4.1 Data Scraping

Data di scrape dari ulasan play store pada aplikasi Line Messenger dengan jumlah data Berupa 7000 data, berbahasa Indonesia dan diambil dari play store dengan region Indonesia. Data memiliki 4 kolom yang berisi :

1. username : Nama Pengguna yang memberikan ulasan
2. content : teks ulasan yang diberikan pengguna
3. score : rating (Bintang) yang diberikan oleh pengguna untuk aplikasi line messenger
4. at : tanggal, bulan, tahun dan jam disaat pengguna memberikan ulasan

4.2 Preprocessing

4.2.1 Data Cleaning

Data dibersihkan dengan menghapus karakter khusus, angka, url, dan spasi berlebih untuk memastikan minimalnya noise pada data saat training model. Data dibersihkan menggunakan regex seperti `r'^[a-zA-Z0-9\s]'` dan `r'\s+'`,

Before	After
"Wah seriusan aplikasi LINE ini keren bgt sih!!! Gue udah pake lama dan selalu puas. mengobrol sama keluarga jd lbh mudah, bisa kirim foto & video juga gampang bgt. Stiker nya lucu2 bingitzzz!!! Terus fitur panggilan video nya jernih, gk pernah error. Pokoknya mantep deh, 5/5 bintang ★★★★"	"Wah seriusan aplikasi LINE ini keren bgt sih Gue udah pake lama dan selalu puas mengobrol sama keluarga jd lbh mudah bisa kirim foto video juga gampang bgt Stiker nya lucu bingitzz Terus fitur panggilan video nya jernih gk pernah error Pokoknya mantep deh bintang"

Tabel 1.1 Data Cleaning

penggunaan regex memudahkan proses data cleaning terkecuali pembersihan emoji yang lebih efisien secara penulisan kode untuk menggunakan library emoji dibanding regex

4.2.2 Case Folding

Tujuan adanya tahap case folding yaitu untuk menyama ratakan semua teks agar model dapat memahami teks lebih mudah tanpa perlu mempertimbangkan atau membedakan antara teks non kapital dan kapital.

Before	After
"Wah seriusan aplikasi LINE ini keren bgt sih Gue udah pake lama dan selalu puas mengobrol sama keluarga jd lbh mudah bisa kirim foto video juga gampang bgt Stiker nya lucu bingitzz Terus fitur panggilan video nya jernih gk pernah error Pokoknya mantep deh bintang"	"wah seriusan aplikasi line ini keren bgt sih gue udah pake lama dan selalu puas mengobrol sama keluarga jd lbh mudah bisa kirim foto video juga gampang bgt stiker nya lucu bingitzz terus fitur panggilan video nya jernih gk pernah error pokoknya mantep deh bintang"

Tabel 1.2 Case Folding

Case Folding tidak perlu menggunakan library external apapun karena adanya modul yang tersedia untuk case folding dari python

4.2.2 Normalization

Tahapan ini dikhkususkan pada alur pembuatan pipeline LDA, setelah data cleaning dan case folding, normalization digunakan untuk menyederhanakan beragam slang, singkatan yang bermakna sama untuk menjadi satu kata yang sama. Tahapan ini menormalisasikan beragam slang dari kamus Bahasa alay dan singkatan yang diambil dari github dengan contoh kata-kata seperti jelek, buruk, jlk, akan di normalisasi menjadi satu kata seperti buruk, buruk, buruk.

Before	After
"wah seriusan aplikasi line ini keren bgt sih gue udah pake lama dan selalu puas mengobrol sama keluarga jd lbh mudah bisa kirim foto video juga gampang bgt stiker nya lucu bingitzz terus fitur panggilan video nya jernih gk pernah error pokoknya mantep deh bintang"	"wah sungguhan aplikasi line ini keren banget sih saya sudah pakai lama dan selalu puas mengobrol sama keluarga jadi lebih mudah bisa kirim foto video juga gampang banget stiker nya lucu banget terus fitur panggilan video nya jernih tidak pernah error pokoknya mantap deh bintang"

Tabel 1.3 Normalization

Kata-kata “alay” seperti Kocak, gokil, ngakak juga berhasil di normalisasikan ke bentuk yang normal digunakan seperti lucu, gila, ketawa. Namun pada ulasan minim ditemukan slang pada ulasan ataupun jenis slang yang masuk kedalam kamus alay dari Github

4.2.3 Stopword

Tahapan ini dikhkususkan pada alur pembuatan pipeline LDA, Setelah normalisasi, stopword digunakan untuk menghapus kata-kata yang dianggap tidak bermakna untuk ekstraksi aspek, sama seperti normalization, stopword menggunakan kamus stopword yang diambil dari github dengan beberapa tambahan kata seperti 'line', 'nya', 'sih', 'yg', 'dah', 'ga', 'gak', 'aja', 'saja', 'juga', 'untuk' menyesuaikan dengan konten teks pada dataset. Stopword berhasil mereduksi kata pada kalimat dari 11 kata menjadi 5-6 kata

Before	After
"wah sungguhan aplikasi line ini keren banget sih saya sudah pakai lama dan selalu puas mengobrol sama keluarga jadi lebih mudah bisa kirim foto video juga gampang banget stiker nya lucu banget terus fitur panggilan video nya jernih tidak pernah error pokoknya mantap deh bintang"	"sungguhan aplikasi line keren banget saya pakai lama selalu puas mengobrol keluarga mudah kirim foto video gampang banget stiker lucu banget fitur panggilan video jernih pernah error pokoknya mantap bintang"

Tabel 1.4 Stopword Removal

4.2.4 Stemming

Tahapan ini dikhkususkan pada alur pembuatan pipeline LDA, Setelah Stopword removal dilakukan stemming untuk menyederhanakan kata agar kembali pada bentuk dasarnya tanpa imbuhan, misalnya memancing menjadi pancing. Stemming sendiri menggunakan kamus bawaan dari sastrawi, tanpa ada kamus tambahan, tetapi di dataset kamus stemming sastrawi mampu menyederhanakan kata-kata pada dataset untuk kembali ke bentuk awal kata.

Before	After
--------	-------

"sungguhan aplikasi line keren banget saya pakai lama selalu puas mengobrol keluarga mudah kirim foto video gampang banget stiker lucu banget fitur panggilan video jernih pernah error pokoknya mantap bintang"	"sungguh aplikasi line keren banget saya pakai lama selalu puas obrol keluarga mudah kirim foto video gampang banget stiker lucu banget fitur panggil video jernih pernah error pokok mantap bintang"
---	--

Tabel 1.5 Stemming

4.2.5 Tokenization

Tokenization terpisah menjadi dua, tokenization pada pipeline LDA dan Indobert. Pada pipeline lda tokenization digunakan untuk memisahkan kata kata pada kalimat seperti “line sangat membantu” menjadi “line”, “sangat”, “membantu” sementara pada indoBERT tokenization dilakukan dengan tambahan per suku kata seperti “Tidak bisa verifikasi nomor telepon” menjadi “tidak bisa”, “verifikasi”, “nomor telepon” karena ini tokenization pada indobert memiliki sedikit kesalahan pada beberapa kalimat, dua kata tidak berhubungan dijadikan satu suku kata.

Before	After
"sungguh aplikasi line keren banget saya pakai lama selalu puas obrol keluarga mudah kirim foto video gampang banget stiker lucu banget fitur panggil video jernih pernah error pokoknya mantap bintang"	["sungguh", "aplikasi", "line", "keren", "banget", "saya", "pakai", "lama", "selalu", "puas", "obrol", "keluarga", "mudah", "kirim", "foto", "video", "gampang", "banget", "stiker", "lucu", "banget", "fitur", "panggil", "video", "jernih", "pernah", "error", "pokok", "mantap", "bintang"]

Tabel 1.6 Tokenization

4.2.5 TF-IDF

Setelah tokenisasi tiap kata, kata kemudian dihitung per bobot kepentingan setiap token menggunakan TF-IDF. Proses Ini bertujuan untuk mengenali kata yang memiliki frekuensi tinggi dalam dokumen tertentu dan juga untuk membedakan antara satu teks ulasan dengan teks lainnya

Before	After

["sungguh", "aplikasi", "line", "keren", "banget", "saya", "pakai", "lama", "selalu", "puas", "obrol", "keluarga", "mudah", "kirim", "foto", "video", "gampang", "banget", "stiker", "lucu", "banget", "fitur", "panggil", "video", "jernih", "pernah", "error", "pokok", "mantap", "bintang"]	{"sungguh": 0.150, "aplikasi": 0.150, "line": 0.150, "keren": 0.150, "banget": 0.450, "saya": 0.150, "pakai": 0.150, "lama": 0.150, "selalu": 0.150, "puas": 0.150, "obrol": 0.150, "keluarga": 0.150, "mudah": 0.150, "kirim": 0.150, "foto": 0.150, "video": 0.300, "gampang": 0.150, "stiker": 0.150, "lucu": 0.150, "banget": 0.150, "panggil": 0.150, "jernih": 0.150, "pernah": 0.150, "error": 0.150, "pokok": 0.150, "mantap": 0.150, "bintang": 0.150}
--	---

Tabel 1.7 Hasil TF-IDF

4.3 Ekstraksi Topik Dan Evaluasi Model LDA

Sebelum melakukan ekstraksi, dilakukan percobaan untuk menentukan jumlah ideal topik dengan meng evaluasi nilai coherence juga perplexity pada tiap model berdasarkan jumlah total topik 20 dan dinilai per 2 topik. Dari hasil ditemukan

Jumlah Topik	Coherence	Perplexity
2	0.0196	1082.88
4	0.0740	1224.87
6	0.1095	1397.21
8	0.1163	1554.16
10	0.1331	1687.17
12	0.1717	1906.61
14	0.1705	1954.61
16	0.1840	2063.32
18	0.1832	2195.84
20	0.1833	2391.79

Tabel 1.8 Hasil Percobaan LDA

Dari hasil evaluasi coherence dan perplexity ditemukan jumlah ideal topik berada pada 16 topik dengan coherence score sejumlah 0.1840 lalu kemudian dilakukan evaluasi lagi berdasarkan keyword untuk menentukan apakah topik bisa dipersempit lagi. Dari hasil evaluasi ditemukan :

Jumlah Topik	Keyword
1	mantap, update, kak, bintang, coba, kasih, kasih bintang, status, admin, suruh
2	suka, suka galat, suka aplikasi, messages, ribettt, sebal, pergi, galat, aplikasi suka, pindah
3	jelek, bantu, aplikasi jelek, akses, senang, aplikasi, fungsi, bad, temu, atur
4	voom, kembali, cerita, buka, fitur, kembali voom, ada, kembali fitur, chat, tidak
5	aneh, blokir, punya, komunikasi, orang, this, aman, jadi, cant, jadi orang
6	pesan, kirim, sampah, iklan, keren, aplikasi, terima, aplikasi sampah, kirim pesan, pemberitahuan
7	masuk, akun, susah, ribet, handphone, aplikasi, pakai, verifikasi, kak, masuk akun
8	baik, bodoh, sandi, lambat, pakai, koneksi, lupa, timeline, aplikasi, lupa sandi
9	tinggal, lumayan, selamat, niat, linenya, bangkrut, media, selamat tinggal, semangat, sosial
10	bagus, oke, good, buruk, aplikasi, virus, aplikasi bagus, terimakasih, mudah, aplikasi buruk
11	user, ganggu, coba baik, iklan, coba, henti, orang, update, depan, yahoo
12	cocok, bisnis, susah, daftar susah, masuk lupa, lupa sandi, sandi, suruh, ganti, lupa
13	galat, voom, fitur, hapus, hilang, kecewa, cerita, voom hilang, voom hapus, galat terus
14	facebook, simpan, foto, masuk facebook, hilang, masuk, video, kenang, kembali, data
15	kocak, parah, versi, kesini, gila, kaia, aplikasi, wallet, aplikasi kocak, konyol
16	nomor, telepon, daftar, nomor telepon, tambah, telpon, nomer, teman, negara, indonesia

Tabel 1.9 Hasil Percobaan LDA

Dari tabel diatas masih ditemukan beberapa kata yang masih bisa dijadikan satu topik seperti “voom hilang” pada topik 13 dan “kembali voom” pada topik 4, aplikasi dan update juga dapat ditemukan di beberapa topik yang berbeda. Dari sini penulis memperkecil topik berdasarkan keyword dan memutuskan untuk memperkecil menjadi 3 topik. Topik – topik tersebut adalah :

- a. Topik 1 (Penggunaan Umum LINE):** Keyword Topic 1 berisi : “bagus, oke, galat, aplikasi, blokir, mantap, teman, terus, pesan, masuk, suka, kirim, terima, google, akun”. Kemudian penulis juga melakukan evaluasi melalui wordcloud yang berisi :



Gambar 1.3 Wordcloud Topik 1

Dari wordcloud dan top keyword, dapat dilihat bahwa topik 1 dominan Membahas tentang penggunaan umum line dan pengalaman umum pengguna, yang berarti topik ini berfokus pada fitur dasar Line contohnya mengirim pesan, menambahkan teman dan notifikasi tidak hanya itu kata seperti bagus, oke, galat juga aplikasi sepertinya membahas tentang kualitas aplikasi secara umum sehingga bisa disimpulkan topik 1 membahas tentang penggunaan umum line.

- b. **Topik 2 (Fitur Tambahan):** Keyword Topic 2 berisi “voom, fitur, hilang, telepon, cerita, good, hapus, nomor, kecewa, jelek, kembali, indonesia, buka, simpan, update”. Kemudian penulis juga melakukan evaluasi melalui wordcloud yang berisi :



Gambar 1.4 Wordcloud Topik 2

Dari top keyword dan wordcloud ditemukan bahwa topik 2 dominan membahas tentang fitur tambahan pada line yang masih ada ataupun yang sudah dihapus di wilayah Indonesia, dibuktikan dengan keyword teratas berupa “voom, hilang dan fitur” bisa disimpulkan bahwa topik 2 membahas tentang fitur tambahan.

- c. **Topik 3 (Login dan Registrasi Akun):** Keyword Topic 3 berisi “masuk, akun, pakai, aplikasi, handphone, kak, daftar, facebook, susah, ribet, nomor, email, buruk, telpon, baik”. Kemudian penulis juga melakukan evaluasi melalui wordcloud yang berisi :



Gambar 1.5 Wordcloud Topik 3

Dari top keyword dan wordcloud ditemukan bahwa topik 3 dominan membahas tentang pendapat pengguna terhadap sistem login, registrasi dan verifikasi akun pada aplikasi. Kesimpulan ini dapat dibuktikan dengan beberapa keyword pada wordcloud seperti “masuk, akun, daftar, email, kode dan verifikasi” bisa disimpulkan bahwa topik 3 membahas tentang Login dan Registrasi Akun

4.5 Modelling Indobert

4.5.1 Komponen Utama Arsitektur Model

- a. **Multi-Head Attention Mechanism:** Model memiliki multi-head attention yang memungkinkan model untuk memahami hubungan kontekstual antar kata dalam kalimat secara bidirectional.
 - b. **Feed-Forward Neural Networks** Setelah lapisan attention, terdapat feed-forward neural network yang memproses representasi yang telah diencode. Lapisan ini menggunakan fungsi aktivasi GELU (Gaussian Error Linear Unit).
 - c. **Layer Normalization dan Residual Connections** Untuk menjaga stabilitas model saat training dan mencegah *vanishing gradient*, model juga memiliki layer normalization dan residual connections pada setiap sub-layer.

4.5.2 Spesifikasi Teknis Model

Berdasarkan implementasi dalam penelitian ini, model IndoBERT yang digunakan memiliki spesifikasi sebagai berikut:

- A. **Model Base:** IndoBERT-base-uncased
- B. **Jumlah Layer:** 12 transformer layers
- C. **Hidden Dimension:** 768
- D. **Attention Heads:** 12 heads per layer
- E. **Vocabulary Size:** ±32,000 tokens
- F. **Maximum Sequence Length:** 512 tokens
- G. **Parameter Count:** ±110 juta parameter

4.5.3 Multi-Topic Classification

Dalam penelitian ini dibuat 3 model terpisah untuk tiap topik dengan arsitektur yang sama namun parameter yang dilatih secara terpisah. Hal ini dilakukan untuk menangani teks yang memiliki topik lebih dari 1.

4.5.4 Tokenization dan Preprocessing

IndoBERT menggunakan algoritma WordPiece tokenization yang memecah kata menjadi sub-word units. Hal ini memungkinkan model untuk menangani kata-kata yang tidak ada dalam vocabulary (out-of-vocabulary words) dan meningkatkan kemampuan generalisasi. Tokens yang digunakan dalam model ini ada 4 yang terdiri dari :

- A. **[CLS]:** Token klasifikasi yang ditambahkan di awal setiap sequence
- B. **[SEP]:** Token separator untuk memisahkan sequence
- C. **[PAD]:** Token padding untuk membuat sequence dengan panjang yang sama
- D. **[UNK]:** Token untuk kata yang tidak dikenal

4.5.5 Fine-Tuning Strategy

Transfer Learning Approach Penelitian ini menggunakan pendekatan transfer learning dimana model IndoBERT yang telah di-pretrain pada korpus bahasa Indonesia besar kemudian di-fine-tune pada dataset yang telah di split dengan Ratio 7:1:2 (train: validation: testing). Setiap model topik dilatih dengan konfigurasi sebagai berikut:

Hyperparameters:

- A. **Learning Rate:** 2e-5
- B. **Batch Size:** 16 (training), 32 (evaluation)
- C. **Epochs:** 3
- D. **Warmup Steps:** 100-500 (bervariasi per topik)
- E. **Weight Decay:** 0.01
- F. **Optimizer:** AdamW

Training Strategy:

- A. **Evaluation Strategy:** Per epoch
- B. **Save Strategy:** Per epoch
- C. **Early Stopping:** Berdasarkan F1-score
- D. **Best Model Selection:** Model dengan F1-score tertinggi

Lalu, untuk memastikan performa model optimal, setiap model menggunakan :

- A. **Load Best Model at End:** True
- B. **Metric for Best Model:** F1-score
- C. **Greater is Better:** True

4.5.6 Evaluasi Model IndoBERT

Setelah topik diidentifikasi, model IndoBERT diterapkan untuk mengklasifikasikan teks sesuai topik tersebut. Hasil evaluasi kinerja model berdasarkan data latih (train) dan data uji (test) adalah sebagai berikut:

TOPIK 1				
Split	Coherence	Precision	Recall	F1-Score
Train : Validation : Test (7:1:2)				
Train	1.2699	0.7805	0.7768	0.7768
Test	1.2064	0.6976	0.6927	0.6917

Tabel 1.10 Evaluasi Model Topik 1

TOPIK 2				
Split	Coherence	Precision	Recall	F1-Score
Train : Validation : Test (7:1:2)				
Train	1.2410	0.6988	0.6711	0.5997
Test	1.2429	0.7027	0.6823	0.6114

Tabel 1.10 Evaluasi Model Topik 2

TOPIK 3				
Split	Coherence	Precision	Recall	F1-Score
Train : Validation : Test (7:1:2)				
Train	1.2410	0.6988	0.6711	0.5997
Test	1.2429	0.7027	0.6823	0.6114

Tabel 1.10 Evaluasi Model Topik 3

Secara keseluruhan, hasil evaluasi menunjukkan model IndoBERT mampu memberikan hasil analisis yang baik jika dievaluasi melalui metrik karena sedikit selisih antara skor pada split data train dan testing. Kebanyakan skor juga berada diatas 0.5 yang menjadi indikasi model berpotensi memberikan hasil yang baik saat dilakukan pengujian dengan data baru, Namun hal ini hanya bisa disimpulkan setelah pengujian.

4.6 Deployment & Uji Model Dengan Data Nyata Baru

Deployment dilakukan pada streamlit untuk web application dan huggingface untuk deployment model prediksi sentimen dan preprocessing.

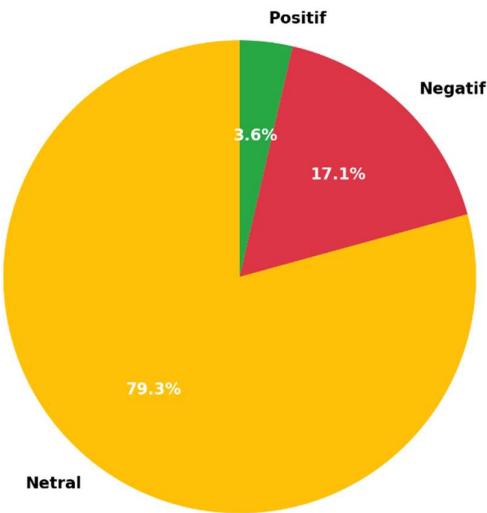
Kemudian dibuat App.py untuk memanggil model prediksi dan preprocessing pada repository huggingface namun pada model preprocessing digunakan fallback berupa pipeline preprocessing lokal pada app.py, demi menghindari terlalu banyak dependency pre processing cukup dilakukan menggunakan menggunakan regex karena minimnya preprocessing yang diperlukan untuk IndoBERT.

Lalu dilakukan tes prediksi dengan nyata yang di scrape dari google play store dengan menggunakan module *google-play-scrapers*, fungsi scraping juga telah terintegrasi pada script app.py. Dari hasil prediksi ditemukan



Gambar 1.6 Hasil uji sentimen per topik

Distribusi Sentimen Keseluruhan



Gambar 1.9 Distribusi hasil uji Sentimen Keseluruhan

Topik	Positif	Netral	Netral
kenapa ga ada layanan people nearby sekarang	Negatif	Netral	Netral
suda apdet malah hilang story nya	Negatif	Netral	Netral
kenapa jadi susah banget untuk masuk ke akun sendiri lewat	Netral	Netral	Netral
ngk bisa di donlod	Netral	Netral	Netral

Gambar 2.0 Tabel Hasil Analisis

4.7 Pembahasan

Dari ekstraksi aspek menggunakan LDA ditemukan 3 topik yang berhubungan dengan review pada line messenger 3 topik tersebut meliputi :

1. Pengalaman umum Penggunaan Line.
2. Fitur Tambahan.
3. Login Dan Registrasi Akun.

Sentimen pada tiap topik cenderung negative dan lebih banyak review yang membahas topik 1. Tidak ada review yang membahas topik 2 sehingga menyebabkan sentimen topik 2 netral sepenuhnya.

Dengan penjelasan untuk tiap sentimen sebagai berikut :

- a. Negatif : ulasan mengkritik atau memberikan keluhan tentang topik pada ulasan
- b. Positif : ulasan memberikan pujian tentang topik tersebut pada ulasan
- c. Netral : ulasan tidak membahas topik tersebut

Pada topik 1 terdapat persebaran yang lumayan seimbang antara Netral dan negatif, namun karena netral bisa berarti review tidak membicarakan topik tersebut bisa diasumsikan bahwa topik 1 dominan negative. Topik 2 netral secara keseluruhan yang berarti 100 data yang di scrape tidak membahas topik 2. Topik 3 dominan netral yang berarti hanya sedikit yang membicarakan topik 3 dan Ketika ada review yang membicarakan topik 3, review tersebut cenderung negative.

Meski hasil evaluasi model IndoBERT > 0.5 namun pada hasil prediksi model cenderung memberikan hasil yang keliru, tidak hanya itu model juga biased terhadap hasil prediksi dengan Biased kepada sentiment netral pada topik 3 dan 2 lalu biased kepada negatif dan netral pada topik 1. Class imbalance dapat menjadi salah satu faktor adanya bias pada prediksi sentiment

```
Distribusi Topic 1 (Pengalaman Umum Penggunaan LINE):
Topic_1_Pengalaman_Umum_Penggunaan_LINE
Negatif      432
Netral       412
Positif      116
Name: count, dtype: int64

Distribusi Topic 2 (Fitur Tambahan):
Topic_2_Fitur_Tambahan
Netral      780
Negatif     177
Positif      3
Name: count, dtype: int64

Distribusi Topic 3 (Login dan Registrasi Akun):
Topic_3_Login_dan_Registrasi_Akun
Netral      605
Negatif     354
Positif      1
Name: count, dtype: int64
Sebelum drop: (999, 5)
Sesudah drop: (960, 5)
```

Gambar 1.9 Class Imbalance Pada Data

Namun, pada saat ini penulis belum menemukan pendekatan yang cocok dan sesuai untuk menanggani class imbalance pada konteks Aspect Based Sentiment Analysis. Tidak hanya itu untuk fine-tuning hanya 950 data sehingga kurangnya variasi representasi dataset yang dapat menjadi salah satu faktor kenapa prediksi model belum optimal.

BAB V

PENUTUP

5.1 Kesimpulan

1. Ada 3 aspek yang berhasil dianalisis menggunakan metode Machine Learning LDA pada review Line Messenger, aspek tersebut adalah :
 - a. Pengalaman penggunaan umum di line
 - b. Fitur tambahan
 - c. Login Dan Registrasi Akun
2. Sentimen pengguna terhadap Line Messenger yang dianalisis menggunakan machine learning memiliki distribusi persentasi sebagai berikut :
 - a. Netral : 79.3%
 - b. Negatif : 17.1%
 - c. Positif : 3.6%
3. Sentimen pengguna terhadap masing masing aspek yang dianalisis menggunakan machine learning memiliki distribusi persentasi sebagai berikut :
 1. Topik 1 :
 - a. Netral : 44.0%
 - b. Negatif : 42.0%
 - c. Positif : 14.0%
 2. Topik 2 :
 - a. Netral : 100%
 - b. Negatif : 0%
 - c. Positif : 0%
 3. Topik 3:
 - a. Netral : 89.0%
 - b. Negatif : 11.0%
 - c. Positif : 0%

Sentiment netral berarti ulasan tidak mengandung topik tersebut dan bukan berarti ulasan memiliki teks yang bersentimen netral Dimana ulasan tidak negative sepenuhnya juga tidak positif sepenuhnya, Disini makin terlihat adanya bias terhadap sentiment netral yang disebabkan oleh class imbalance

3. Performa model indoBERT untuk tiap topik sebagai berikut :

- a. **Topik 1:** Precision model pada data latih sekitar 0,780 dan recall 0,776 dengan F1 sekitar 0,776. Pada data train, presisi menurun menjadi ~0,697, recall ~0,692, dan F1 ~0,691. Ini menunjukkan model cukup akurat mengenali ulasan yang berkaitan dengan penggunaan umum LINE pada data latih, tetapi kinerjanya menurun pada data test meski demikian penurunan kinerja tidak signifikan sehingga tidak ada indikasi overfit atau underfit.
- b. **Topik 2:** Pada data latih, presisi ~0,698 dan recall ~0,671 menghasilkan F1 ~0,599. Pada data uji, presisi ~0,702, recall ~0,682, dan F1 ~0,611. Topik ini berbeda dengan Topik 1 yang memiliki penurunan kinerja pada data test, melainkan penurunan terjadi peningkatan skor pada data test tetapi tidak tidak signifikan sehingga tidak ada indikasi overfit atau underfit.
- c. **Topik 3:** Data latih menunjukkan presisi ~0,698, recall ~0,671, dan F1 ~0,599. Pada data uji, presisi ~0,702, recall ~0,682, dan F1 ~0,611. Performa untuk topik ini relatif serupa dengan topik 2 tanpa ada indikasi underfit atau overfit.

5.2 Saran

Dataset masih sangat imbalanced sehingga perlu penerapan class balancing agar hasil optimal jika class imbalanced model akan cenderung bias pada kelas mayoritas sehingga output model tidak optimal dengan model yang tidak optimal maka pelaku industry akan cenderung melakukan kesalahan dalam mengambil Keputusan bisnis berdasarkan hasil Analisa sehingga diperlukan class balancing atau cara alternatif lain untuk menanggulangi class imbalance agar model tidak biased dalam melakukan klasifikasi , Jumlah data pada dataset juga perlu lebih banyak saat ingin melakukan fine tuning pada indobert untuk memperbanyak variasi representasi pada dataset ini juga dapat digunakan sebagai cara alternatif menanggulangi class imbalance pada dataset.

DAFTAR PUSTAKA

- [1] Eck, “Kinerja Line setelah 13 tahun di Tanah Air | Indotelko,” IndoTelko. <https://www.indotelko.com/read/1719249362/di-usia-13-tahun-ini-capaian-line-di-tanah-air#:~:text=dicapai%20di%20Indonesia%20dan%20global,179%20juta%20pengguna%20aktif%20bulanan>
- [2] A. S. Wardani, “500 Juta Stiker Line Dikirim di Indonesia Sepanjang Tahun 2024,” *liputan6.com*, Jan. 05, 2025. [Online]. Available: <https://www.liputan6.com/tekno/read/5862380/500-juta-stiker-line-dikirim-di-indonesia-sepanjang-tahun-2024#:~:text=Liputan6.com%2C%20Jakarta%20,percakapan%20di%20Indonesia%20melalui%20Line>
- [3]. Topik.Id, “4 Layanan LINE resmi ditutup di Indonesia, kalah bersaing?,” *Trending Topik*. <https://www.topik.id/2025/02/4-layanan-line-resmi-ditutup-di-indonesia-kalah-bersaing.html#:~:text=2>
- [4] Ibm, “Analisis Sentimen,” IBM.com, Dec. 26, 2023. <https://www.ibm.com/id-id/topics/sentiment-analysis#:~:text=Analisis%20sentimen%20menggunakan%20teknologi%20pemrosesan,menangani%20skenario%20yang%20lebih%20kompleks>
- [5] H. Imaduddin, F. Y. A’la, and Y. S. Nugroho, “Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 8, Jan. 2023, doi: 10.14569/ijacsa.2023.0140813.
- [6] K. D. Indarwati and H. Februariyanti, “Analisis sentimen terhadap kualitas pelayanan aplikasi GoJek menggunakan metode Naive Bayes Classifier,” *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, vol. 10, no. 1, Mar. 2023, doi:10.35957/jatisi.v10i1.2643.
- [7] N. Nurhasiyah, R. Dwiyansaputra, S. I. Murpratiwi, and A. Aranta, “Analisis Sentimen Pengguna Platform Media Sosial X pada Topik Pemilihan Presiden 2024 Menggunakan Perbandingan Model Monolingual dan Multilingual BERT,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 1, pp. 626–634, Dec. 2024, doi:10.36040/jati.v9i1.12430.
- [8] D. S. Putri, N. Sulistiyowati, and A. Voutama, “Analisis Sentimen dan Pemodelan Ulasan Aplikasi AdaKami Menggunakan Algoritma SVM dan KNN,” *Journal Sensi*, vol. 9, no. 2, pp. 209–225, Aug. 2023, doi:10.33050/sensi.v9i2.2914
- [9] Y. A. Singgalen, “Performance analysis of IndoBERT for sentiment classification in Indonesian hotel review data,” *ejurnal.seminar-id.com*, Jan. 2025, doi:10.47065/josh.v6i2.6505
- [10] F. A. D. Aryanti, A. Luthfiarta, and D. A. I. Soeroso, “Aspect-Based Sentiment Analysis with LDA and IndoBERT Algorithm on Mental Health App: Riliv,” *Jurnal AI and Computing*, vol. 9, no. 2, Mar. 2025, doi:10.30871/jaic.v9i2.8958.

- [11] J. R. Jim, M. A. R. Talukder, P. Malakar, M. M. Kabir, K. Nur, and M. F. Mridha, “Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review,” *Natural Language Processing Journal*, vol. 6, p. 100059, Feb. 2024, doi: 10.1016/j.nlp.2024.100059.
- [12] Y. C. Hua, P. Denny, J. Wicker, and K. Taskova, “A systematic review of aspect-based sentiment analysis: domains, methods, and trends,” *Artificial Intelligence Review*, vol. 57, no. 11, Sep. 2024, doi: 10.1007/s10462-024-10906-z.
- [13] A. Jazuli, N. Widowati, and R. Kusumaningrum, “Aspect-based sentiment analysis on student reviews using the Indo-Bert base model,” *E3S Web of Conferences*, vol. 448, p. 02004, Jan. 2023, doi: 10.1051/e3sconf/202344802004.
- [14] N. M. Gardazi, A. Daud, M. K. Malik, A. Bukhari, T. Alsahfi, and B. Alshemaimri, “BERT applications in natural language processing: a review,” *Artificial Intelligence Review*, vol. 58, no. 6, Mar. 2025, doi: 10.1007/s10462-025-11162-5.
- [15] E. Dave and A. Chowanda, “IPerFEX-2023: Indonesian personal financial entity extraction using indoBERT-BiGRU-CRF model,” *Journal of Big Data*, vol. 11, no. 1, Sep. 2024, doi: 10.1186/s40537-024-00987-6.
- [16] S. Dami and R. Alimardani, “An Aspect-Level sentiment analysis based on LDA topic modeling,” *Journal of Information Systems and Telecommunication (JIST)*, vol. 12, no. 46, pp. 117–126, Jun. 2024, doi: 10.61186/jist.38104.12.46.117.
- [17] J. Chavan, “NLP: Tokenization , stemming , lemmatization , bag of Words ,TF-IDF , POS,” Medium, Dec. 14, 2021. [Online]. Available: <https://medium.com/@jeevanchavan143/nlp-tokenization-stemming-lemmatization-bag-of-words-tf-idf-pos-7650f83c60be>
- [18] N. Yuyun, A. D. Latief, T. Sampurno, N. Hazriani, A. O. Arisha, and N. Mushaf, “Next Sentence Prediction: The Impact of Preprocessing Techniques in Deep Learning,” *Next Sentence Prediction: The Impact of Preprocessing Techniques in Deep Learning*, pp. 274–278, Oct. 2023, doi: 10.1109/ic3ina60834.2023.10285805.
- [19] Bustamin, A, Prayogi, AA, Siswanto, D, Rafrin, M and Nurdin, A 2025, “TEXT NORMALIZATION FOR INDONESIAN SLANG WORDS IN SENTIMENT ANALYSIS DEVELOPMENT”, *ICIC Express Letters, Part B: Applications*, vol. 16, no. 2, pp. 121-129. <https://doi.org/10.24507/icicelb.16.02.121>