

**ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN
APLIKASI LINE MESSENGER MENGGUNAKAN
ALGORITMA INDOBERT & LDA**

SKRIPSI

Diajukan 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**

2026

**ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN
APLIKASI LINE MESSENGER MENGGUNAKAN
ALGORITMA INDOBERT & 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**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN
APLIKASI LINE MESSENGER MENGGUNAKAN
ALGORITMA INDOBERT & LDA**

yang disusun dan diajukan oleh

ATQIYA TRIANDA PUTRA ANUGRAH

22.11.5083

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Januari 2026

Dosen Pembimbing,



Anna Baita, M. Kom.

NIK. 190302290

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN
APLIKASI LINE MESSENGER MENGGUNAKAN
ALGORITMA INDOBERT & LDA

yang disusun dan diajukan oleh

ATQIYA TRIANDA PUTRA ANUGRAH

22.11.5083

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393

Lukman, S.Kom., M.Kom.
NIK. 190302151

Anna Baita, S.Kom., M.Kom.
NIK. 190302290

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Atqiya Trianda Putra Anugrah
NIM : 22.11.5083

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Line Messenger Menggunakan Algoritma Indobert & LDA

Dosen Pembimbing : Anna Baita, M. Kom.

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 Januari 2026

Yang Menyatakan,



Atqiya Trianda Putra Anugrah

HALAMAN PERSEMBAHAN

Puji dan Syukur dipanjatkan kepada Allah SWT yang senantiasa memberikan kekuatan, kesehatan dan kemudahan sehingga skripsi ini dapat diselesaikan. Skripsi ini dipersembahkan oleh penulis kepada :

1. Ibu dan Almarhum ayah penulis juga seluruh keluarga penulis yang memberikan dukungan, doa dan bantuan finansial hingga penelitian ini dapat terselesaikan.
2. Ibu Anna Baita, M.Kom selaku dosen pembimbing yang telah memberikan waktu dan tenaga untuk membimbing penulis dari awal hingga akhir penelitian dan penyusunan skripsi.
3. Seluruh dosen Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta yang telah membagikan beragam ilmu pengetahuan dan wawasan selama masa studi penulis di Universitas AMIKOM Yogyakarta.

Penulis berharap skripsi ini dapat memberikan manfaat bagi pengguna aplikasi *messaging* khususnya aplikasi LINE serta memberikan manfaat untuk pengembangan ilmu pengetahuan pada bidang *data mining* dan *deep learning*

KATA PENGANTAR

Puji dan Syukur dipanjatkan kepada Allah SWT atas segala Rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi *LINE* Messenger Menggunakan Algoritma IndoBERT & LDA”. Skripsi ini disusun sebagai syarat penyelesaian masa studi untuk memperoleh gelar sarjana pada program studi informatika. Selama penyusunan skripsi, penulis mendapatkan berbagai bantuan, dan dukungan dari berbagai pihak. Oleh karena itu penulis ingin menyampaikan terima kasih sebesar-besarnya kepada :

1. Seluruh pihak keluarga penulis yang selalu memberikan bantuan dan dukungan kepada penulis dari awal hingga akhir masa studi
2. Ibu Anna Baita, M.Kom selaku dosen pembimbing atas bimbingan dan arahan yang dibagikan kepada penulis selama penyusunan skripsi
3. Seluruh tim dosen penguji atas kritik dan saran yang diberikan untuk meningkatkan skripsi ini
4. Seluruh dosen Universitas AMIKOM Yogyakarta atas segala ilmu dan pengalaman yang dibagikan kepada penulis selama masa studi
5. Alfarion Raihan Putra, selaku saudara dari ibu yang berbeda atas segala dukungan, doa dan bantuan dari awal hingga akhir masa studi.
6. Seluruh teman-teman penulis serta pihak lain yang telah membantu penulis, yang tidak dapat disebutkan satu per satu.
7. Seluruh pengguna *LINE Messenger* yang telah menulis ulasan di *google play store* sehingga penulis dapat menganalisis ulasan-ulasan tersebut

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan dan keterbatasan sehingga kritik dan saran sangat diharapkan demi peningkatan di masa yang mendatang. Penulis berharap skripsi ini memberikan manfaat baik untuk berbagai pihak dalam mengetahui pendapat pengguna serta pengembangan ilmu pengetahuan *deep learning* bagi praktisi *data mining*

Yogyakarta, 17 Januari 2026

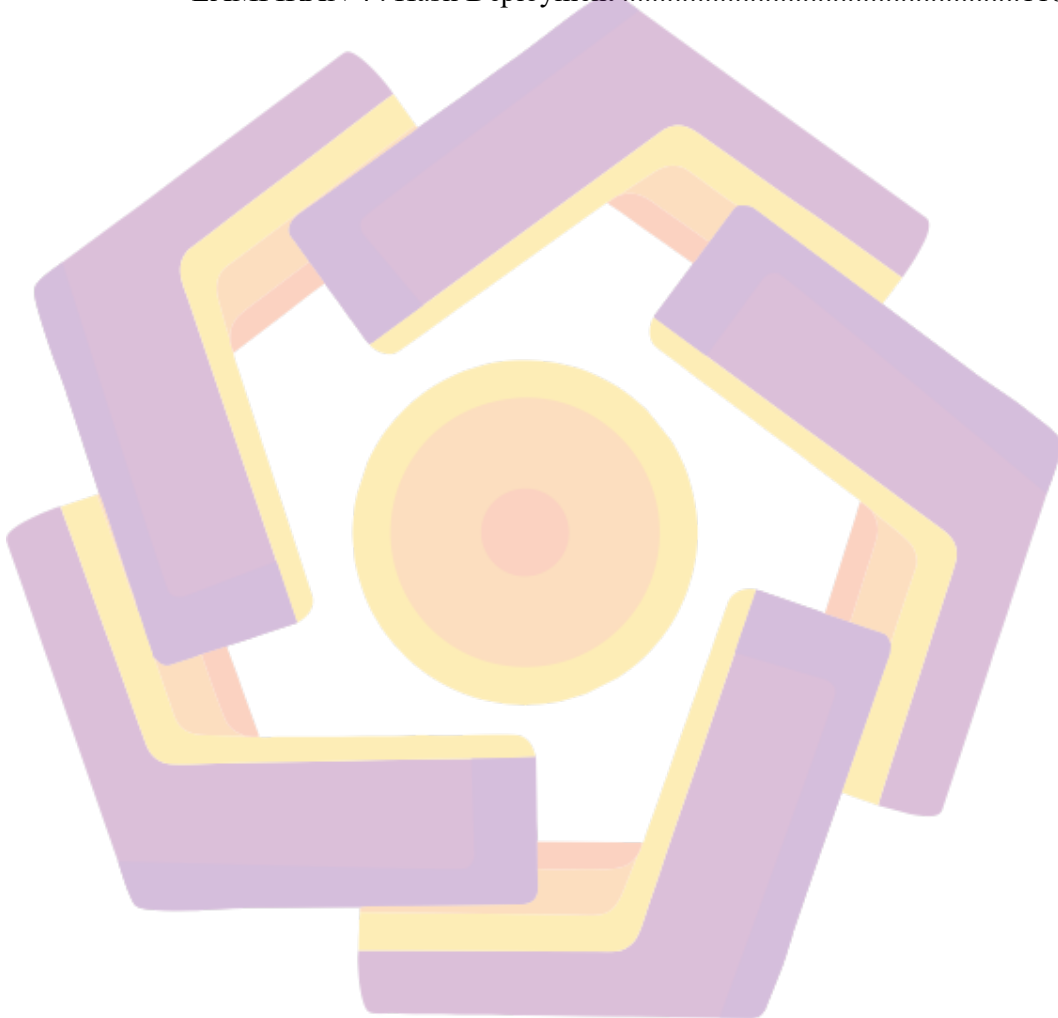
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xiv
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori.....	29
2.2.1 <i>LINE Messenger</i>	29
2.2.2 Analisis Sentimen	29
2.2.3 Analisis Sentimen Berbasis Aspek	30
2.2.4 <i>Supervised Learning</i>	31
2.2.5 <i>Unsupervised Learning</i>	31
2.2.6 <i>Pre-Processing</i>	31
2.2.6.1 <i>Data Cleaning</i>	32
2.2.6.2 <i>Case Folding</i>	32
2.2.6.3 <i>Normalization</i>	32

2.2.6.4	<i>Stopword Removal</i>	33
2.2.6.5	<i>Stemming</i>	33
2.2.6.6	<i>Tokenizing</i>	34
2.2.7	<i>Feature Extraction With Sublinear TF-IDF</i>	35
2.2.8	LDA	36
2.2.9	<i>Data Labeling</i>	38
2.2.10	BERT	39
2.2.11	IndoBERT	43
2.2.13	<i>Fine-Tuning</i>	44
2.2.14	Model Evaluation	44
BAB III METODE PENELITIAN		47
3.1	Objek Penelitian	47
3.2	Alur Penelitian	48
3.3	Alat dan Bahan	51
BAB IV HASIL DAN PEMBAHASAN		53
4.1	Data Scraping	53
4.2	<i>Preprocessing</i>	54
4.2.1	<i>Data Cleaning dan Case Folding</i>	54
4.2.2	<i>Tokenizing</i>	55
4.2.3	<i>Normalization</i>	55
4.2.4	<i>Stopword Removal</i>	57
4.2.5	<i>Stemming</i>	57
4.3	<i>Feature Extraction</i>	58
4.4	Stage One	59
4.4.1	Aspect Extraction dan Aspect Labeling	59
4.5	Data Labeling	62
4.6	Stage Two	63
4.6.1	EDA	63
4.6.2	Data Split	65
4.6.3	Sentiment Classification And Model Evaluation With Indobert	67
4.6.4	Deployment	78
BAB V PENUTUP		80
5.1	Kesimpulan	80
5.2	Saran	81

REFERENSI	82
LAMPIRAN.....	91
LAMPIRAN 1 : EDA	91
LAMPIRAN 2 : Rincian Komputasi Dan Data Split.....	93
LAMPIRAN 3 : Script Deployment	100
LAMPIRAN 4 : Hasil Deployment	118



DAFTAR TABEL

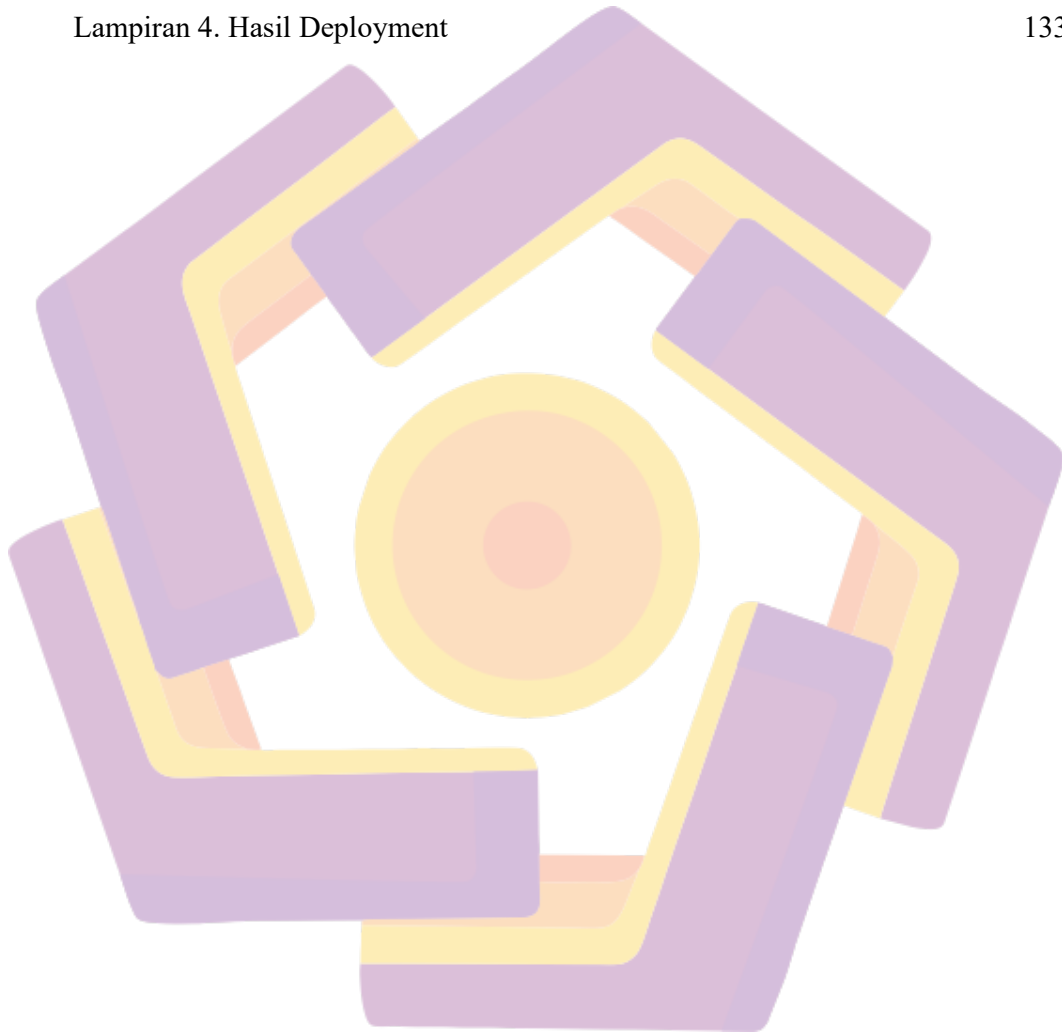
Tabel 2.1 Keaslian Penelitian	24
Tabel 2.2 Arsitektur IndoBERT	60
Tabel 2.3 Confusion Matrix	62
Tabel 4.1 Sampel Data	70
Tabel 4.2 <i>Data Cleaning</i> dan <i>Case Folding</i>	71
Tabel 4.3 Tokenization	72
Tabel 4.4 Dictionary	72
Tabel 4.5 Normalization	73
Tabel 4.6 Stopword Removal	74
Tabel 4.7 Stemming	75
Tabel 4.8 Feature Extraction	75
Tabel 4.9 Coherence Score dan Perplexity	77
Tabel 4.10 Data Dengan Label Yang Salah	79
Tabel 4.11 Hasil Validasi Ulang Label	80
Tabel 4.12 Distribusi Sentimen	80
Tabel 4.13 Sampel data dengan token terpanjang	81
Tabel 4.14 Distribusi Token	82
Tabel 4.15 Hasil Model Dengan Variasi Data Splitting	83
Tabel 4.16 <i>Split Data</i> 8:1:1 Pada Topik 3	84
Tabel 4.17 <i>Tokenizer IndoBERT</i>	85
Tabel 4.18 Percobaan <i>Epoch</i> dan <i>Learning Rate</i>	86
Tabel 4.19 <i>Learning Rate</i> Paling Optimal	88
Tabel 4.20 Data Eksternal	90
Tabel 4.21 Data Eksternal Modifikasi	91
Tabel 4.22 Evaluasi Data Baru <i>Epoch 5</i>	92
Tabel 4.23 Evaluasi Data Baru <i>Epoch 2</i>	94

DAFTAR GAMBAR

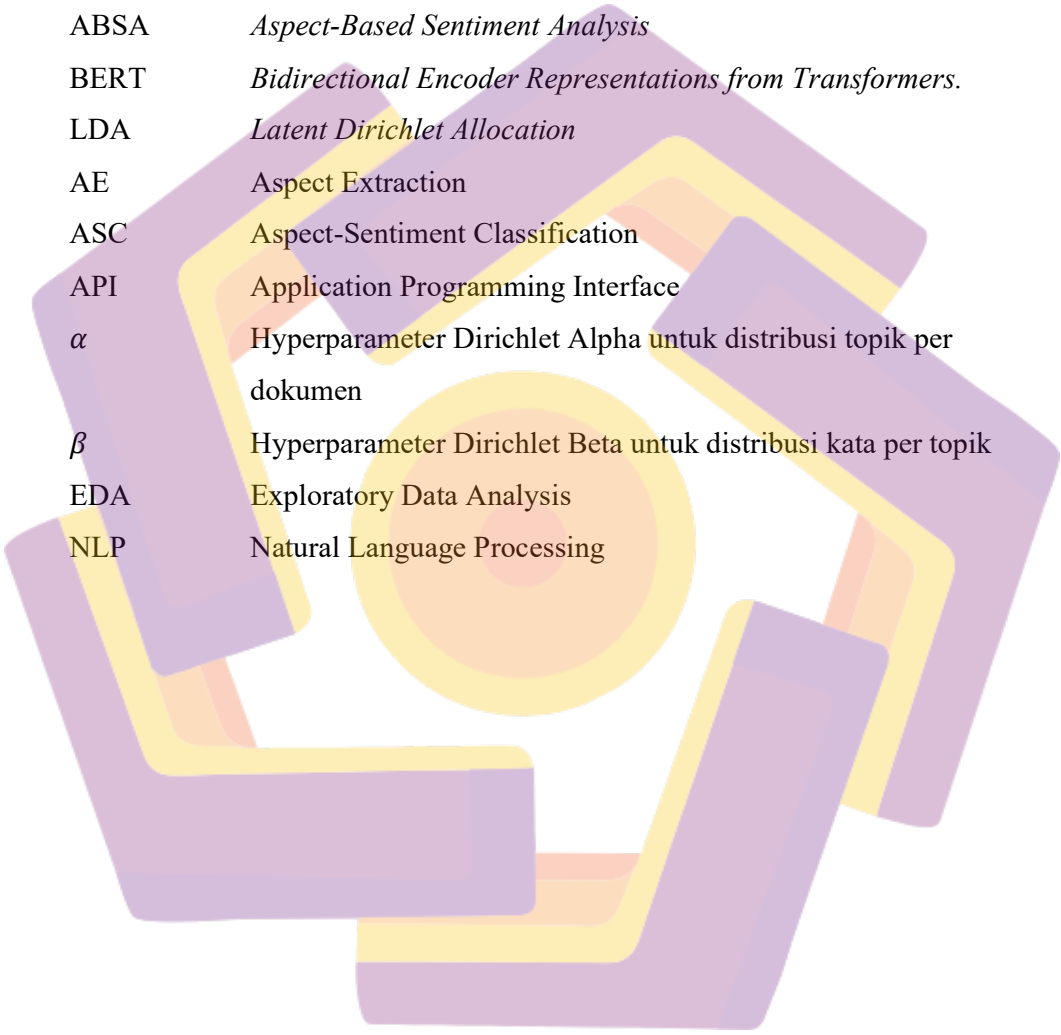
Gambar 2.1. Model Grafis LDA	54
Gambar 2.2 Arsitektur dan <i>Framework</i> BERT	57
Gambar 2.3 Representasi <i>input</i> pada BERT	58
Gambar 3.1 Alur Penelitian	65
Gambar 4.1 <i>Wordcloud</i> 5 Topik	76
Gambar 4.2 <i>Wordcloud</i> 3 Topik	77
Gambar 4.3 <i>Deployment</i>	96
Gambar 4.4 Statistik Pada <i>Deployment</i>	96

DAFTAR LAMPIRAN

Lampiran 1. EDA	107
Lampiran 2. Rincian Komputasi Dan Data Split	108
Lampiran 3. Script Deployment	115
Lampiran 4. Hasil Deployment	133



DAFTAR LAMBANG DAN SINGKATAN



SVM	Support Vector Machines
TF-IDF	<i>Term Frequency – Inverse Document Frequency</i>
NMF	<i>Non-negative Matrix Factorization</i>
LSA	<i>Latent Semantic Analysis</i>
ABSA	<i>Aspect-Based Sentiment Analysis</i>
BERT	<i>Bidirectional Encoder Representations from Transformers.</i>
LDA	<i>Latent Dirichlet Allocation</i>
AE	Aspect Extraction
ASC	Aspect-Sentiment Classification
API	Application Programming Interface
α	Hyperparameter Dirichlet Alpha untuk distribusi topik per dokumen
β	Hyperparameter Dirichlet Beta untuk distribusi kata per topik
EDA	Exploratory Data Analysis
NLP	Natural Language Processing

DAFTAR ISTILAH

Aspect Extraction

Subtugas ABSA untuk menemukan aspek/topik yang dibahas dalam dokumen.

Aspect-Sentiment Classification

Subtugas ABSA untuk menentukan polaritas (pos/neg/netral) pada tiap aspek yang diekstraksi.

Topic Modeling

Metode unsupervised untuk menemukan topik

WordPiece

Algoritma tokenisasi sub-kata yang digunakan oleh model BERT/IndoBERT.

IndoBERT

Varian BERT monolingual untuk Bahasa Indonesia.

INTISARI

LINE Messenger yang diluncurkan pada 2011 mengalami penurunan jumlah pengguna aktif dari 300 juta menjadi 196 juta pada 2023. Ulasan pengguna menjadi sumber informasi penting untuk mengetahui sentimen pengguna terhadap penggunaan aplikasi LINE dan aspek apa saja yang dibahas dalam ulasan aplikasi LINE. Untuk mencapai tujuan tersebut perlu menggunakan metode *Aspect-Based Sentiment Analysis* (ABSA). Penelitian ini juga bertujuan menguji performa IndoBERT dalam menganalisis ulasan LINE Messenger berbahasa Indonesia.

Untuk mengekstraksi topik digunakan LDA yang menghasilkan 3 topik berdasarkan kejelasan topik (kata kata dominan pada wordcloud) dan kecocokan pada konteks LINE Messenger. tiga topik tersebut adalah “Pengalaman Umum Penggunaan LINE”, “Fitur Tambahan” dan “Login Dan Registrasi”.

Untuk klasifikasi sentimen digunakan algoritma IndoBERT dengan performa terbaik pada epoch lima dengan rentang F1-Score 0.88, 0.94 dan 0.96. Namun Epoch 2 mempunyai waktu komputasi lebih singkat dengan performa yang masih kompetitif. Mayoritas sentimen bersifat negatif dengan total 4.788 sentimen negatif dari 5.278 ulasan.

Ini menunjukkan bahwa ABSA menggunakan LDA dan IndoBERT dapat memberikan *insight* penting tentang pendapat pengguna dan topik apa saja yang dibahas oleh pengguna dalam ulasan LINE. Sentimen yang negatif menandakan adanya ketidakpuasan pengguna terhadap layanan LINE sehingga penelitian ABSA diperlukan untuk menganalisis informasi ini lebih dalam.

Kata kunci: LINE, ABSA, BERT, IndoBERT, LDA

ABSTRACT

LINE Messenger, which was launched in 2011, experienced a decline in the number of active users from 300 million to 196 million in 2023. User reviews have become an important source of information for understanding user sentiment toward the use of the LINE application and identifying which aspects are discussed in LINE app reviews. To achieve this objective, the Aspect-Based Sentiment Analysis (ABSA) method is employed. This study also aims to evaluate the performance of IndoBERT in analyzing Indonesian-language LINE Messenger reviews.

Topic extraction is performed using Latent Dirichlet Allocation (LDA), which produces three topics based on topic clarity (dominant words visualized in word clouds) and their relevance to the LINE Messenger context. The three identified topics are “General User Experience of LINE,” “Additional Features,” and “Login and Registration.”

For sentiment classification, the IndoBERT algorithm is used, achieving its best performance at the fifth epoch with F1-scores of 0.88, 0.94, and 0.96. However, the second epoch offers a shorter computation time while maintaining competitive performance. The majority of the sentiments are negative, with a total of 4,788 negative sentiments out of 5,278 reviews.

These results indicate that ABSA using LDA and IndoBERT can provide valuable insights into user opinions and the main topics discussed in LINE app reviews. The predominance of negative sentiment reflects user dissatisfaction with LINE’s services, highlighting the importance of applying ABSA to analyze this information in greater depth.

Keyword: LINE, ABSA, BERT, IndoBERT, LDA