

Analisis Sentimen Berbasis Aspek pada Ulasan Aplikasi Gojek

R A Rahman¹, V H Pranatawijaya², N N K Sari^{*3}

¹⁻³Program Studi Teknik Informatika, Universitas Palangka Raya

E-mail: reshananda@mhs.eng.upr.ac.id¹, viktorhp@it.upr.ac.id²,
ovanoorks@it.upr.ac.id^{*3}

Abstrak. Penggunaan aplikasi *mobile* meningkat pesat di era digital, termasuk Gojek, aplikasi populer di Indonesia yang menyediakan layanan transportasi, pesan antar makanan, dan pembayaran digital. Ulasan pengguna di Play Store menunjukkan berbagai masalah yang memerlukan perhatian. Ulasan ini memberikan wawasan tentang pandangan pengguna, memungkinkan identifikasi masalah, dan pengembangan layanan. Dengan teknik *Aspect Based Sentiment Analysis* (ABSA), pandangan pengguna dapat dipahami lebih baik, membantu evaluasi dan perbaikan aplikasi Gojek untuk meningkatkan kualitas layanan dan kepuasan pengguna. Penelitian ini bertujuan menganalisis sentimen berdasarkan aspek-aspek dalam ulasan pengguna aplikasi Gojek di Play Store dalam bahasa Inggris, dengan mencari pola sentimen yang akurat dan mengidentifikasi aspek yang perlu diperbaiki. Data diambil dari ulasan pengguna aplikasi Gojek di Google Play Store. Teknik pemodelan topik *Latent Dirichlet Allocation* (LDA) digunakan untuk mengidentifikasi topik-topik relevan. Pelabelan sentimen dilakukan menggunakan model BERT, sementara evaluasi sentimen dan aspek dilakukan dengan model *distilbert-base-uncased-finetuned-sst-2-english*. Hasil menunjukkan bahwa model BERT mencapai akurasi tertinggi untuk sentimen sebesar 96.67% dan aspek *Service* sebesar 98.78%. Terdapat ruang untuk perbaikan terutama pada aspek *user experience*, *service*, dan *payment*. Faktor-faktor yang mempengaruhi akurasi termasuk distribusi sentimen, jumlah data, *preprocessing*, dan model yang digunakan.

Kata kunci: Gojek; ulasan pengguna; *Aspect Based Sentiment Analysis* (ABSA); sentimen; akurasi

Abstract. *Mobile app usage is increasing rapidly in the digital era, including Gojek, a popular app in Indonesia that provides transportation, food delivery, and digital payment services. User reviews in the Play Store indicate various issues that require attention. These reviews provide insight into user views, enabling problem identification and service development. With the Aspect Based Sentiment Analysis (ABSA) technique, user views can be better understood, helping evaluate and improve the Gojek application to improve service quality and user satisfaction. This research aims to analyze sentiment based on aspects of user reviews of the Gojek application on the Play Store in English by finding accurate sentiment patterns and identifying aspects that need to be improved. The data was taken from user reviews of the Gojek application on the Google Play Store. Latent Dirichlet Allocation (LDA) topic modeling technique was used to identify relevant topics. Sentiment labeling was performed using the BERT model, while sentiment and aspect evaluation were performed with the distilbert-base-uncased-finetuned-sst-2-english model. The results showed that the BERT model achieves the highest accuracy for sentiment at 96.67% and Service aspects at 98.78%. There is room for improvement, especially in the user experience, service, and payment*

aspects. Factors affecting accuracy include sentiment distribution, amount of data, preprocessing, and the model used.

Keywords: Gojek; user reviews; Aspect Based Sentiment Analysis (ABSA); sentiment; accuracy

1. Pendahuluan

Di era digital saat ini, aplikasi *mobile* telah menjadi bagian integral dari kehidupan sehari-hari, memfasilitasi berbagai kebutuhan pengguna dengan cepat dan efisien. Di Indonesia, salah satu *platform* multi-layanan terkemuka adalah Gojek, yang menyediakan layanan transportasi daring, pesan antar makanan, pembayaran digital, dan berbagai layanan lainnya. Dengan lebih dari 100 juta unduhan dan *rating* 4,6 di Play Store, Gojek telah menjadi pilihan utama bagi banyak pengguna di Asia Tenggara.

Namun, di balik angka-angka tersebut, terdapat informasi berharga dari ulasan pengguna yang tersebar di platform Play Store. Ulasan tersebut mencerminkan pengalaman langsung pengguna dalam menggunakan aplikasi Gojek, baik dari sisi positif maupun negatif. Dengan lebih dari 5,78 juta ulasan, data ini menjadi sumber potensial untuk memahami pandangan pengguna dan mengidentifikasi area-area yang perlu perbaikan.

Untuk menggali lebih dalam pandangan pengguna, penelitian ini menggunakan pendekatan *Aspect Based Sentiment Analysis* (ABSA). ABSA adalah model yang dapat mengekstraksi aspek-aspek tertentu dari sebuah ulasan dan menilai sentimen yang terkait dengan masing-masing aspek tersebut [1]. Karena ABSA dapat mengekstraksi aspek-aspek tertentu dari sebuah ulasan dan menilai sentimen yang terkait dengan masing-masing aspek tersebut, sehingga memungkinkan kita untuk memahami opini pengguna terkait berbagai aspek aplikasi Gojek. Dengan demikian, penelitian ini bertujuan untuk menganalisis ulasan pengguna aplikasi Gojek untuk memahami opini mereka terkait berbagai aspek aplikasi.

Beberapa penelitian terdahulu telah memfokuskan pada analisis sentimen dan klasifikasi aspek menggunakan berbagai algoritma dalam konteks beragam domain. Dalam penelitian oleh [1], mereka menerapkan algoritma *Random Forest* (RF) untuk klasifikasi aspek pada data set bahasa Indonesia, dengan hasil terbaik diperoleh pada domain restoran dengan *precision* 79,3%, *recall* 91,3%, dan *F1-Score* 84,3%. Selanjutnya, penelitian oleh [2] membandingkan kinerja algoritma *Naïve Bayes* (NB), *Support Vector Machine* (SVM), dan k-NN dalam analisis sentimen terhadap komentar tentang *Gadget Samsung Galaxy Z Flip 3* di YouTube. SVM menunjukkan hasil terbaik dengan rata-rata akurasi sebesar 96,43%, dengan aspek desain mencapai 94,40% dan aspek harga mencapai 97,44%.

Penelitian terdahulu telah mengeksplorasi analisis sentimen dan klasifikasi aspek dalam berbagai konteks. Misalnya, dalam penelitian [3], evaluasi sentimen terhadap aplikasi KAI Access menggunakan *Naive Bayes Classifier* (NBC) dan SVM mencapai akurasi rata-rata 91,63% dan *F1-Score* 75,55% setelah *tuning hyperparameter*. Sementara itu, penelitian [4] yang memfokuskan pada deteksi emosi dan sentimen pada tweet terkait COVID-19 menunjukkan bahwa klasifikasi berbasis aspek dan korelasi mencapai akurasi tertinggi sebesar 97%. Selain itu, [5] menerapkan analisis sentimen berbasis aspek pada Taman Nasional Bromo Tengger Semeru menggunakan berbagai model *machine learning*, dengan transfer *learning* dari IndoBERT memberikan hasil terbaik dengan akurasi 91,48% dan *F1-Score* 71,56%.

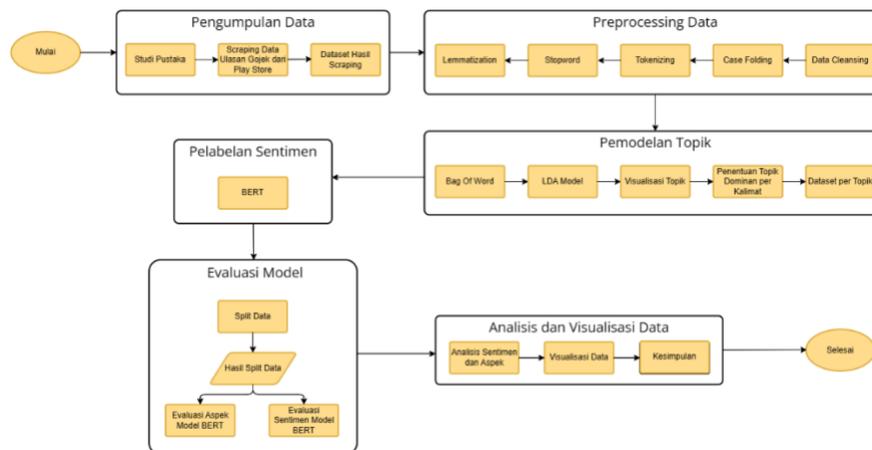
Melalui teknik pemodelan topik *Latent Dirichlet Allocation* (LDA), penulis akan mengidentifikasi topik-topik yang relevan dalam ulasan pengguna. Untuk analisis sentimen, penulis akan menggunakan berbagai metode, termasuk BERT. Evaluasi sentimen dan aspek akan dilakukan menggunakan model *Support Vector Machine* (SVM) dan *Bidirectional Encoder Representations from Transformers* (BERT). LDA adalah metode yang efektif dalam menemukan kelompok topik tersembunyi dalam dokumen dengan pendekatan probabilistik yang mudah diinterpretasikan, sementara BERT unggul dalam representasi kontekstual dan performa tinggi pada berbagai tugas NLP berkat kemampuannya menangkap nuansa dan makna teks secara mendalam. Kombinasi LDA dan BERT memungkinkan penulis mendapatkan pemahaman yang komprehensif mengenai topik dan sentimen ulasan pengguna, di mana LDA mengidentifikasi topik utama dan BERT mengevaluasi sentimen secara mendalam. Dengan tambahan

evaluasi menggunakan SVM, pendekatan ini memberikan hasil analisis yang kaya, akurat, dan terperinci, sehingga memberikan wawasan yang mendalam dari data ulasan pengguna.

Tujuan utama dari penelitian ini adalah untuk menemukan pola sentimen dengan akurasi yang tinggi dan mengidentifikasi aspek-aspek tertentu yang perlu diperbaiki dalam aplikasi Gojek. Dengan demikian, penelitian ini diharapkan dapat memberikan wawasan berharga bagi pengembang aplikasi untuk meningkatkan pengalaman pengguna dan menjaga kepuasan pengguna dalam menggunakan layanan Gojek.

2. Metode

Tahapan dari metode yang diusulkan ditunjukkan pada Gambar 1. Proses dimulai dari pengumpulan data, *preprocessing* data, pemodelan topik, pelabelan sentimen, evaluasi model, dan analisis dan evaluasi data.



Gambar 1. Alur Metode Penelitian

2.1 Pengumpulan Data

Pada tahapan ini penulis mengumpulkan data terkait jurnal maupun dari *website* yang relevan dengan penelitian terkait analisis sentimen berbasis aspek, BERT, dan metode lain yang dipakai pada penelitian ini. *Web scraping* adalah proses pengambilan data dari halaman web untuk analisis lebih lanjut [6]. *Web scraping* yaitu teknik di mana program komputer mengekstrak data dari *output* yang dapat dibaca manusia yang berasal dari program lain [7]. Teknik pengambilan data atau *scraping* data dengan menggunakan *library google-play-scrapers* yang tersedia dalam bahasa pemrograman *Python*.

2.2 Preprocessing Data

Pre-processing merupakan langkah penting sebelum pemodelan data, yang melibatkan persiapan teks dan penghapusan informasi tidak relevan untuk memastikan kebersihan dan kemudahan pengolahan data. Dalam penelitian ini, digunakan lima tahap *preprocessing* teks: pembersihan data, *case folding*, *tokenizing*, penghapusan *stopwords*, dan *lemmatization* [8]. Tahap awal melibatkan pembersihan teks dengan menghapus karakter tidak relevan seperti nama pengguna, spasi kosong, karakter khusus, dan URL [9]. Dilanjutkan dengan *case folding* untuk mengonversi teks menjadi huruf kecil dan menghapus tanda baca yang tidak relevan [1]. Tahap berikutnya adalah *tokenizing*, yaitu memisahkan kalimat menjadi urutan kata-kata untuk proses selanjutnya [10], diikuti dengan penghapusan *stopwords* untuk mempercepat analisis [8]. Terakhir, *lemmatization* digunakan untuk mengurangi kata berinfleksi menjadi bentuk dasarnya [11].

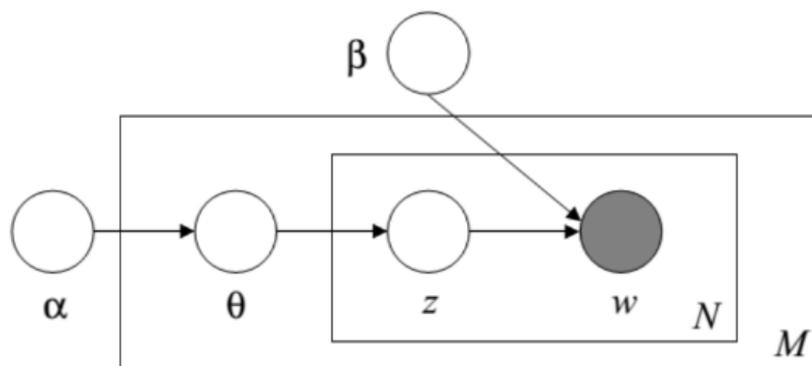
2.3 Pemodelan Topik

Pada tahap pemodelan topik, data set hasil pra-pemrosesan teks direpresentasikan sebagai vektor numerik menggunakan algoritma *Bag of Words* (BoW). BoW merupakan representasi kata-kata unik dari korpus teks ulasan, diekstraksi dengan mengidentifikasi kata-kata berdasarkan *Parts of Speech* (PoS), dan direpresentasikan sebagai vektor [12]. Selanjutnya, vektor tersebut dimasukkan ke dalam variabel *dictionary* untuk diproses menggunakan algoritma *Latent Dirichlet Allocation* (LDA). LDA merupakan metode *unsupervised learning* yang digunakan untuk mengelompokkan data menjadi beberapa kelas dan merangkum serta memproses data yang berukuran besar [13]. Proses ini melibatkan beberapa iterasi untuk membentuk model LDA dan mengevaluasi koherensinya untuk menentukan jumlah topik yang paling sesuai [14]. Setelah model LDA terbentuk, dilakukan visualisasi topik untuk mempermudah interpretasi, dan data set dibagi ke dalam masing-masing topik yang telah ditentukan sebelumnya. Hal ini dilakukan dengan menambahkan interpretasi topik berdasarkan nilai dominan topik yang diperoleh dari LDA. Selanjutnya, nilai dominan topik dipetakan ke dalam aspek yang telah ditetapkan sebelumnya. Pada akhirnya, dibuat sebuah data set yang berisi ulasan untuk setiap topik yang telah ditetapkan.

Model LDA ditunjukkan pada gambar di bawah ini tentang generatif proses pada LDA dijelaskan sebagai berikut [15]:

1. Pilih $N \sim \text{Poisson}(x)$,
2. Pilih $\theta \sim \text{Dir}(\alpha)$,
3. Untuk setiap N kata W_n ,
 - a. Pilih Topik $Z_n \sim \text{Multinomial}(\theta)$,
 - b. Pilih sebuah kata W_n dari $p(W_n | Z_n, \beta)$

Bentuk dari representasi model LDA dapat digambarkan dalam diagram pada Gambar 2, yang memberikan visualisasi yang jelas tentang hubungan antara topik-topik yang dihasilkan dan kata-kata yang mewakilinya.



Gambar 2. Representasi Model *Latent Dirichlet Allocation*

2.4 Pelabelan Sentimen

BERT digunakan dalam penelitian ini untuk evaluasi sentimen dan aspek pada data teks yang tidak berlabel. BERT adalah model arsitektur *transformer* yang dipakai untuk Pemrosesan Bahasa Alami (NLP) dan Analisis Sentimen, dilatih pada korpus teks tanpa label [12] [16].

2.5 Model Evaluasi Sentimen dan Aspek

Sebelum melakukan evaluasi dengan menggunakan metode BERT, pertama-tama dilakukan ekstraksi fitur dengan menggunakan model *distilbert-base-uncased-finetuned-sst-2-english*. Model BERT tersebut menggunakan parameter *learning rate* sebesar $1e-5$, *batch size* 32, *max length* 128, dan *train epoch* 3, serta

proses ini berjalan pada GPU. Proses ekstraksi fitur ini melibatkan penggunaan *tokenizer* BERT untuk menghasilkan token dari data ulasan. Untuk menentukan panjang kalimat *tokenized* maksimum, data ulasan ditandai dengan metode '*encode*' dari *tokenizer* BERT, dan kemudian dilihat kalimat terpendek [12]. Setelah proses ekstraksi fitur selesai, dilakukan evaluasi model menggunakan metode BERT untuk analisis yang lebih mendalam terhadap data ulasan.

Selain itu, dalam penelitian ini, dilakukan penggunaan metrik akurasi dan *F1-Score* untuk mengukur evaluasi kinerja setiap pengklasifikasian pada data set Ulasan aplikasi Gojek. Akurasi didefinisikan sebagai tingkat kedekatan antara nilai yang diprediksi dengan nilai aktual. Sementara itu, *F1-Score* merupakan evaluasi komprehensif dari rata-rata tertimbang *precision* dan *recall*. *Precision* adalah nilai ketepatan antara informasi *ground truth* dengan respon yang diputuskan oleh sistem, sedangkan *recall* adalah tingkat keberhasilan untuk mendapatkan kembali informasi oleh sistem. Kedua metrik ini memberikan gambaran yang holistik tentang kinerja pengklasifikasi dalam menangani data set ulasan aplikasi Gojek [17].

2.6 Analisis dan Visualisasi Data

Pada tahap ini, dilakukan analisis sentimen dan aspek pada data set ulasan aplikasi Gojek. Hasil analisis sentimen ini akan menunjukkan jumlah dan akurasi sentimen positif dan negatif pada setiap aspek yang diamati. Analisis ini memberikan wawasan tentang persepsi pengguna terhadap berbagai aspek aplikasi Gojek dan dapat digunakan untuk meningkatkan pengalaman pengguna secara keseluruhan.

3. Hasil dan Pembahasan

3.1 Scraping Data

Pada proses ini dilakukan pengumpulan data dari ulasan aplikasi Gojek di *website* Google Play Store. Proses pengambilan data dilakukan mulai dari bulan Januari 2020 hingga Desember 2023. Jumlah total ulasan yang berhasil diambil dari proses pengambilan data adalah sebanyak 14,147 ulasan. Data terbagi dalam empat kolom: *username*, *content*, *score*, dan *at*. Contoh hasil *scraping* data ulasan aplikasi Gojek pada Tabel 1.

Tabel 1. Hasil *Scraping* Ulasan Aplikasi Gojek

No	UserName	Content	Score	at
1	Obet	<i>It keeps crashing when im about to order. Seems like</i>	2	12/12/2023
	Nyawo	<i>there's a bug? Please fix it asap because it's annoying</i>		4:15
2	rahma	<i>After updating, yesterday i still can use it, now i can't</i>	1	12/12/2023
	omceee	<i>using the apps, when i open just blank and then back</i>		2:47
3		<i>to the phone menu.</i>		
	Any	<i>The app crashed and caused my order to be placed</i>	1	12/14/2023
4		<i>twice. My GoPay was charged 2 times, yet I did not</i>		5:00
	Priam	<i>enter my PIN. This sucks! Fix the app</i>		
5	Sampaio	<i>The new mode does fast-booking is abusive - I don't</i>	1	12/12/2023
		<i>even have the time to hit "return" that gojek booked</i>		0:56
5		<i>me a ride and refuses to show me the price until the</i>		
	Avior	<i>trip is finished.</i>		
5	Jonas	<i>Not good apps, the gojar driver that I booked</i>	3	12/21/2023
		<i>occasionally far away from my position, and we need</i>		23:44
		<i>is good respond for each driver.</i>		

3.2 Preprocessing Data

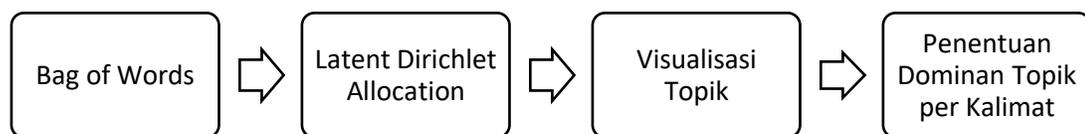
Setelah memperoleh data set kemudian dilakukan beberapa tahapan *preprocessing* yang bertujuan untuk membersihkan dan memproses data teks agar siap untuk dianalisis. Berikut ini tahapan *preprocessing* pada Tabel 2.

Tabel 2. Contoh Hasil *Preprocessing* Data

No	Pemrosesan Awal	Data Mentah	Pembersihan Data
1	Data Cleansing	<i>It keeps crashing when im about to order Seems like theres a bug Please fix it asap because its annoying</i>	<i>It keeps crashing when im about to order Seems like theres a bug Please fix it asap because its annoying</i>
2	Case Folding	<i>It keeps crashing when im about to order Seems like theres a bug Please fix it asap because its annoying</i>	<i>it keeps crashing when im about to order seems like theres a bug please fix it asap because its annoying</i>
3	Tokenizing	<i>it keeps crashing when im about to order seems like theres a bug please fix it asap because its annoying</i>	<i>['it', 'keeps', 'crashing', 'when', 'i', 'm', 'about', 'to', 'order', 'seems', 'like', 'there', 's', 'a', 'bug', 'please', 'fix', 'it', 'asap', 'because', 'its', 'annoying']</i>
4	Stopword Removal	<i>['it', 'keeps', 'crashing', 'when', 'i', 'm', 'about', 'to', 'order', 'seems', 'like', 'there', 's', 'a', 'bug', 'please', 'fix', 'it', 'asap', 'because', 'its', 'annoying']</i>	<i>['keeps', 'crashing', 'order', 'seems', 'like', 'bug', 'please', 'fix', 'asap', 'annoying']</i>
5	Lemmatization	<i>['keeps', 'crashing', 'order', 'seems', 'like', 'bug', 'please', 'fix', 'asap', 'annoying']</i>	<i>keep crashing order seem bug fix</i>

3.3 Pemodelan Topik

Pada tahap pemodelan topik ini, menggunakan algoritma *Latent Dirichlet Allocation* (LDA) sebagai metode pengelompokan topik dari data ulasan aplikasi Gojek di Google Play Store. Pada pembuatan model LDA ini, terdapat alur yang perlu dilakukan, langkah-langkahnya dapat dilihat pada Gambar 3.



Gambar 3. Alur *Topic Modelling*

3.3.1 Bag of Words

Pada tahap pemodelan topik, selain kamus (*dictionary*), *input* yang digunakan adalah representasi vektor dari kamus tersebut. Hasil pembobotan kata akan digunakan oleh model *topic modelling* untuk memetakan dan menganalisis data. Pembobotan kata ini menggunakan pendekatan *Bag of Words* yang telah disediakan dalam modul Gensim.

3.3.2 Latent Dirichlet Allocation Model

Pada tahap ini, setelah melakukan pembobotan *Bag of Words*, kemudian dibuat model *Latent Dirichlet Allocation* (LDA) untuk mengekstrak topik yang terdapat pada ulasan aplikasi Gojek. Model tersebut dievaluasi menggunakan *coherence score* untuk menilai seberapa baik model yang dibuat. Dalam model

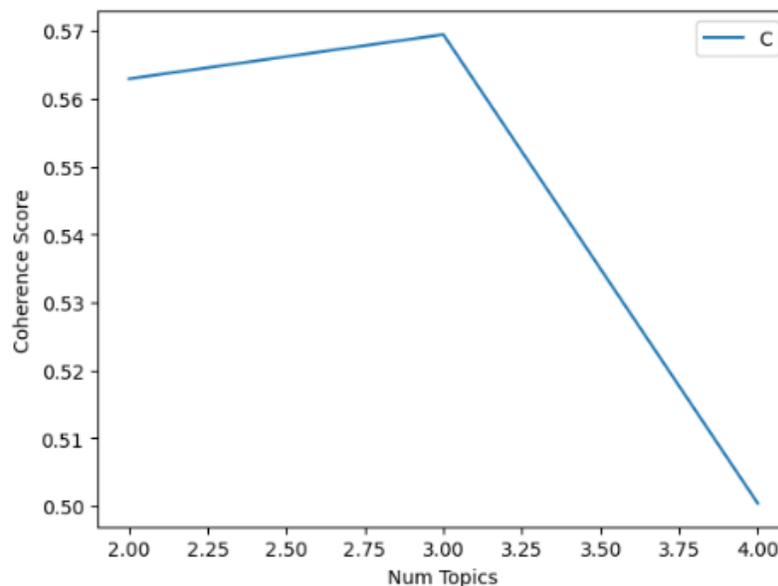
LDA ini, terdapat beberapa parameter utama yang menjadi *input*, yaitu *Bag of Words*, jumlah topik, kamus (*dictionary*), *chunksize*, *passes*, *random state*, dan *alpha*. Parameter "*passes*" digunakan untuk menentukan jumlah iterasi atau pelatihan model. Pada parameter jumlah topik, model LDA akan mengidentifikasi pola topik berdasarkan distribusi kata dalam dokumen dan mencoba menentukan topik yang ada sesuai dengan jumlah topik yang telah ditentukan. Pada proses ini, digunakan *coherence score* untuk mengevaluasi seberapa *interpretable* dan koheren topik-topik yang ditemukan oleh model.

Dari hasil proses model LDA di atas menggunakan limit lima, yang bertujuan untuk membatasi maksimum jumlah topik yang ingin dieksplorasi, serta start dua untuk menentukan jumlah topik minimum yang ingin dieksplorasi, dan step satu sebagai langkah penambahan jumlah topik. Berikut adalah hasil nilai koherensi untuk setiap topik pada Tabel 3.

Tabel 3. Hasil *Coherence Value* Pada Setiap Topik

<i>Num Topics</i>	<i>Coherence Value</i>
2	0.563
3	0.569
4	0.5

Setelah mendapatkan nilai koherensi di atas, langkah selanjutnya adalah menampilkan data tersebut dalam grafik nilai koherensi yaitu jumlah topik 3 yang memiliki *coherence values* tertinggi. Berikut adalah tampilan grafik nilai koherensi pada Gambar 4.

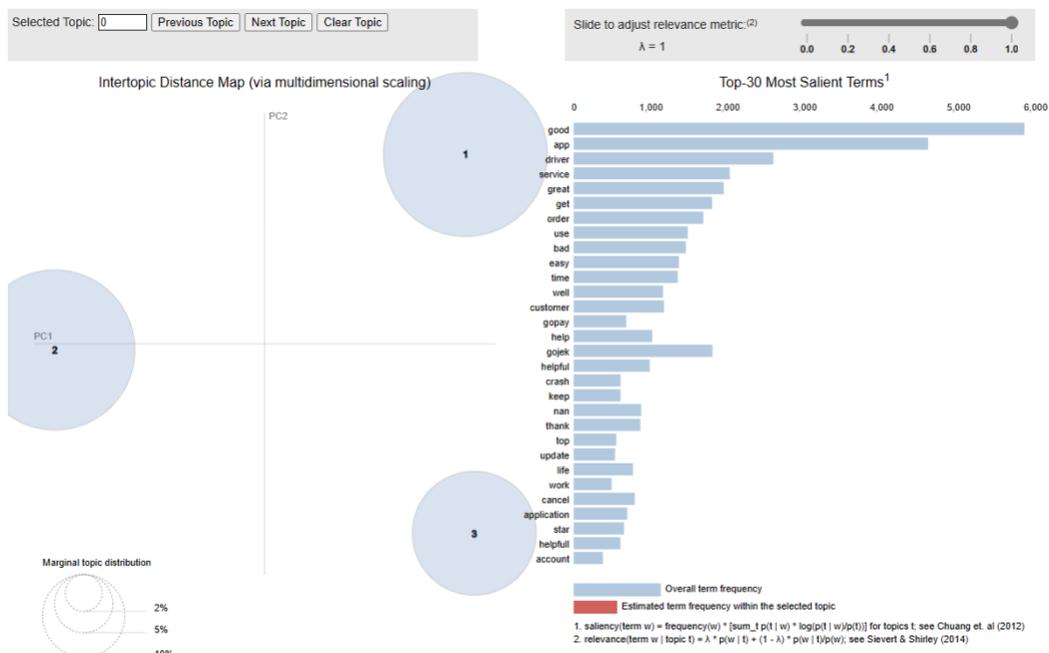


Gambar 4. Grafik Hasil *Coherence Value*

3.3.3 Visualisasi Topik

Pada tahap ini, hasil dari model *Latent Dirichlet Allocation* sebelumnya akan divisualisasikan menggunakan *pyLDAvis*. Dari visualisasi ini, akan diinterpretasikan topik-topik yang ada berdasarkan kata-kata yang telah dilakukan pembobotan dan pemodelan LDA. Visualisasi topik ini akan menggunakan nilai koherensi tertinggi dari pemodelan LDA sebelumnya, yaitu dengan jumlah topik 3. Ini berarti dari data set ulasan yang telah diproses, terdapat 3 topik yang paling sering muncul. Dari proses di atas, maka hasil yang

didapat adalah jumlah topik 3 karena pada tahap pemodelan topik dengan LDA memiliki nilai koherensi tertinggi. Berikut adalah tampilan dari visualisasi topik menggunakan model LDA pada Gambar 5.



Gambar 5. Hasil Visualisasi Topik Model

Berdasarkan visualisasi yang dihasilkan dari interpretasi topik yang didapatkan, kata-kata yang sering muncul pada setiap topik dapat digunakan untuk menentukan aspek-aspek yang relevan pada topik tersebut. Aspek-aspek yang didapatkan dari jumlah topiknya tiga adalah *user experience*, *service*, dan *payment*. Berikut adalah aspek-aspek yang telah ditentukan beserta dengan kata-kata per topik, yang dapat dilihat pada Tabel 4.

Tabel 4. Sebaran Kata Tiap Jumlah Topik

Topik	Aspek	Kata Per Topik
1	<i>User Experience</i>	<i>good, app, service, great, gojek, use, easy, well, help, helpful, nan, thank, life, make, application, helpfull, nice, love, much, cool</i>
2	<i>Service</i>	<i>driver, get, order, bad, time, customer, always, cancel, star, find, give, food, far, ride, take, money, location, gofood, go, know</i>
3	<i>Payment</i>	<i>gopay, crash, keep, top, update, work, still, account, close, fix, new, suck, email, bug, try, change, force, sere, say, payment</i>

3.3.4 Penentuan Dominan Topik Per Kalimat

Pada tahap ini, data set dibagi ke dalam masing-masing topik yang telah ditentukan sebelumnya. Hal ini dilakukan dengan menambahkan interpretasi topik berdasarkan nilai dominan topik yang diperoleh. Setelah mendapatkan nilai dominan topik, langkah selanjutnya adalah memetakan nilai dominan topik ke dalam aspek yang telah ditentukan sebelumnya. Jika nilai dominan topiknya adalah nol, maka kalimat tersebut akan dimasukkan ke dalam aspek *user experience*. Jika nilai dominan topiknya adalah satu, maka kalimat tersebut akan dimasukkan ke dalam aspek *service*. Sedangkan, jika nilai dominan topiknya adalah dua,

maka kalimat tersebut akan dimasukkan ke dalam aspek *payment*. Berikut ini adalah contoh hasil dari penentuan dominan topik dapat dilihat pada Tabel 5.

Tabel 5. Contoh Hasil Penentuan Dominan Topik Per Kalimat

<i>Dominant_topic</i>	<i>Perc_Contrib</i>	<i>Topic_Keywords</i>	<i>Data_Lemmatized</i>	<i>Content</i>
0	0.845	good, app, service, great, gojek, use, easy, well, help, helpful	['really', 'helpfull', 'app', 'become', 'easy', 'simple']	really helpfull app become easy simple
1	0.628	driver, get, order, bad, time, customer, always, cancel, star, find	['location', 'appear', 'map', 'confirm']	location appear map confirm
2	0.552	gopay, crash, keep, top, update, work, still, account, close, fix	['keep', 'crash', 'wrong', 'adress', 'everytime', 'enter']	keep crash wrong adress everytime enter
1	0.598	driver, get, order, bad, time, customer, always, cancel, star, find	['driver', 'rude', 'car', 'old', 'dirty', 'pricing', 'wise', 'also', 'lose', 'grab', 'waste', 'money', 'use', 'grab']	driver rude car old dirty pricing wise also lose grab waste money use grab

3.4 Pelabelan Sentimen

Dalam tahap ini, pelabelan sentimen dilakukan menggunakan BERT dari *library Transformers*. Jenis BERT yang digunakan adalah "*distilbert/distilbert-base-uncased-finetuned-sst-2-english*" yang memiliki *output* untuk sentimen positif dan negatif. Pada tahap pelabelan sentimen menggunakan BERT ini, terdapat beberapa atribut yang perlu ditambahkan.

Pada proses ini dihasilkan data sebanyak 14.037 data, dengan 7.840 data positif dan 6.197 data negatif. Berikut adalah contoh sedikit data yang telah diberi label dapat dilihat pada Tabel 6.

Tabel 6. Pelabelan BERT

Teks	Sentimen	Score
<i>wait long minute taxi much well value book service try</i>	<i>Positive</i>	<i>0.876546</i>
<i>current version crash time new animation introduce use food estimate delivery time accurate anymore</i>	<i>Negative</i>	<i>0.986086</i>
<i>use less get point promos bad system</i>	<i>Negative</i>	<i>0.992186</i>
<i>pay transfer use say phone connect internet internet good balance deduct fix cause pay twice</i>	<i>Negative</i>	<i>0.988188</i>
<i>part daily greatly aid nice addition app include dark mode</i>	<i>Positive</i>	<i>0.994038</i>

3.5 Model Evaluasi Sentimen dan Aspek

Untuk model BERT yang digunakan adalah *distilbert-base-uncased-finetuned-sst-2-english*. Pada model BERT ini, objek *tokenizer* dibuat untuk memecah teks menjadi token-token kecil. Token-token ini kemudian dapat dimasukkan ke dalam model BERT untuk analisis lebih lanjut, seperti klasifikasi teks atau generasi teks. Penggunaan *tokenizer* yang sudah disesuaikan dengan model BERT memastikan bahwa teks yang diberikan ke model akan diproses dengan cara yang sesuai dengan cara model tersebut telah dilatih.

Parameter BERT yang digunakan meliputi *learning rate* (1e-5), *batch size* 32, *max_seq_length* 128, dan *train epoch* 3.

Pada data set “Ulasan Aplikasi Gojek” yang memiliki label aspek dan sentimen. Kemudian dibagi menjadi data latih dan data uji. Berikut ini adalah daftar pembagian data latih dan data uji yang dapat dilihat pada Tabel 7.

Tabel 7. Pembagian Data Ulasan Aplikasi Gojek

Metode	Klasifikasi	Data Latih (85%)	Data Uji (15%)
BERT	<i>Positive</i>	5261	936
	<i>Negative</i>	6670	1170
	Total	11931	2106

3.5.1. Klasifikasi Sentimen

Pada proses ini dihasilkan data hasil evaluasi untuk aspek dari data set "Ulasan Aplikasi Gojek". Berikut adalah data evaluasi terdapat pada Tabel 8.

Tabel 8. Evaluasi Sentimen Model BERT

Data set	Pelabelan	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Ulasan Aplikasi Gojek	BERT	96,67%	97,58%	94,87%	96,20%

Berdasarkan hasil evaluasi sentimen pada Tabel 8, menunjukkan bahwa hasil evaluasi model yang digunakan pada penelitian ini mendapatkan akurasi sebesar 96,67% dan *F1-Score* 96,20%. Pada penelitian [5] mendapatkan nilai akurasi 91,48% dan *F1-Score* 71,56%. Pada penelitian [3] mendapatkan hasil akurasi 91,63% dan *F1-Score* 75,55%. Berdasarkan hasil evaluasi yang telah dipaparkan menunjukkan bahwa model yang digunakan penulis berhasil mendapatkan nilai akurasi tertinggi dibandingkan kedua penelitian tersebut, nilai akurasi menandakan bahwa model yang diusulkan dapat dengan benar mengklasifikasikan sentimen dengan keakuratan yang relatif tinggi.

3.5.2. Klasifikasi Aspek

Pada evaluasi aspek menggunakan model BERT, digunakan data set "Ulasan Aplikasi Gojek". Kedua data set tersebut dievaluasi menggunakan pelabelan aspek dari BERT. Pada proses ini dihasilkan data hasil evaluasi untuk aspek dari data set "Ulasan Aplikasi Gojek". Berikut adalah data evaluasi terdapat pada Tabel 9.

Tabel 9. Evaluasi Aspek Model BERT Data Set Ulasan Aplikasi Gojek

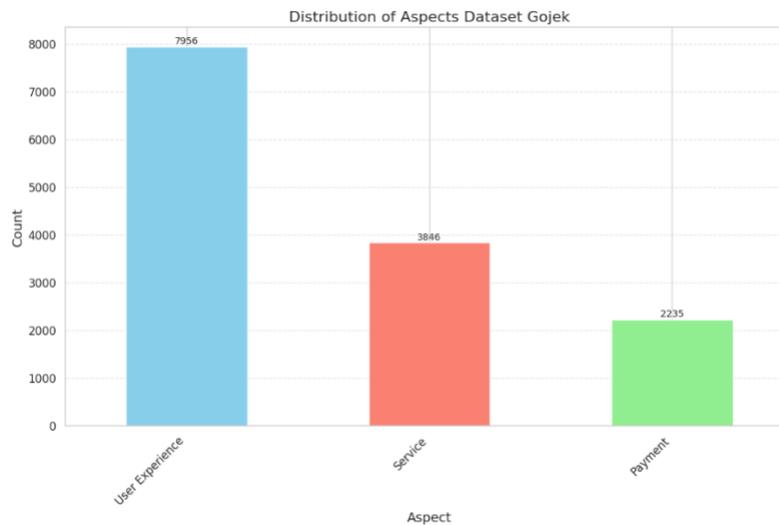
Pelabelan	Aspek	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
BERT	<i>Service</i>	98,78%	96%	94,73%	95,36%
	<i>Payment</i>	96,42%	86,11%	91,57%	83,78%
	<i>User Experience</i>	97,65%	97,94%	98,66%	98,30%

Berdasarkan hasil evaluasi aspek pada Tabel 9, menunjukkan bahwa hasil evaluasi model yang digunakan pada penelitian ini mendapatkan nilai akurasi dari aspek *Service* sebesar 98,78%, aspek *Payment* 96,42%, dan aspek *user experience* 97,65%. Pada penelitian [2] mendapatkan nilai akurasi dari aspek desain sebesar 94,40%, aspek harga sebesar 97,44%, aspek spesifikasi sebesar 96,22%, dan aspek citra merk sebesar 97,63%. Pada penelitian [1] mendapatkan akurasi untuk Nilai *precision* mencapai 79,3%, *recall* 91,3%, dan *F1-Score* 84,3% untuk domain restoran; 93,0% *precision*, 87,0% *recall*, dan 89,3% *F1-Score*

untuk domain *e-commerce*; serta 91,8% *precision*, 81,2% *recall*, dan 85,7% *F1-Score* untuk domain hotel. Pada penelitian [4] mendapatkan nilai akurasi dan *F1-Score* untuk aspek rumah mendapatkan akurasi 97%. Berdasarkan hasil evaluasi yang telah dipaparkan menunjukkan bahwa model yang digunakan penulis berhasil mendapatkan nilai akurasi tertinggi dibandingkan ketiga penelitian tersebut, nilai akurasi menandakan bahwa model yang diusulkan dapat dengan benar mengklasifikasikan sentimen dengan keakuratan yang relatif tinggi.

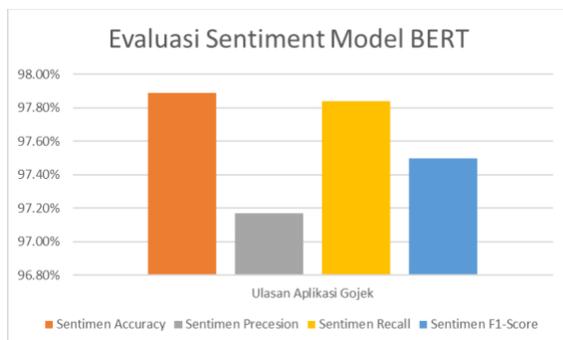
3.6 Analisis dan Visualisasi Data

Pada tahap sebelumnya, aspek-aspek dari ulasan aplikasi Gojek di Play Store ditentukan menggunakan model *Latent Dirichlet Allocation (LDA)*. Berikut ini adalah distribusi dari proses penentuan topik ini dapat dilihat pada Gambar 6.

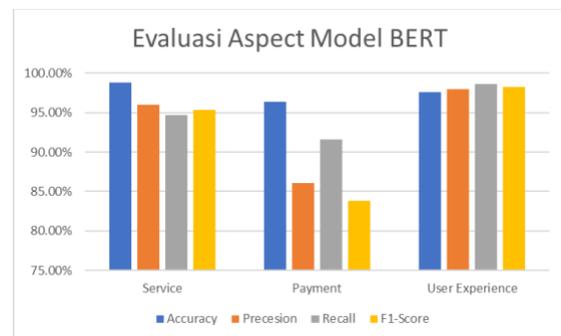


Gambar 6. Distribusi Aspek Data Set Ulasan Gojek

Selanjutnya, hasil evaluasi sentimen pada data set Ulasan Aplikasi Gojek menggunakan model BERT didapatkan hasil evaluasi untuk pelabelan menggunakan BERT memiliki akurasi tertinggi yaitu 96,67%, dengan presisi 97,58%, *recall* 94,87%, dan *F1-Score* 96,20%. Berikut visualisasi evaluasi sentimen pada Gambar 7 dan Gambar 8.



Gambar 7. Evaluasi Sentimen Model BERT



Gambar 8. Evaluasi Aspek Model BERT.

Untuk data set Ulasan Aplikasi Gojek, hasil evaluasi menunjukkan bahwa pelabelan menggunakan *labeling* menggunakan BERT menunjukkan performa tertinggi dengan akurasi 98,78% untuk aspek *Service*, 96,42% untuk aspek *payment*, dan 97,65% untuk aspek *user experience*.

4. Kesimpulan

Penelitian ini menganalisis ulasan aplikasi Gojek dari Januari 2020 hingga Desember 2023 menggunakan model LDA untuk menemukan tiga aspek utama: *user experience*, *service*, dan *payment*. Metode BERT digunakan untuk pelabelan sentimen. Evaluasi model BERT menunjukkan performa yang sangat baik dengan akurasi 96,67%, presisi 97,58%, *recall* 94,87%, dan *F1-Score* 96,20%. Faktor-faktor seperti distribusi sentimen, jumlah data, proses *preprocessing*, dan model evaluasi mempengaruhi analisis sentimen dan aspek.

Pengguna Gojek memberikan ulasan beragam terkait *user experience*, *service*, dan *payment*. Ulasan positif mencatat kemudahan, kecepatan, dan efisiensi aplikasi, serta kepuasan terhadap layanan. Namun, ulasan negatif menyebutkan kesulitan penggunaan, seringnya terjadi kesalahan, dan masalah dengan fitur. Performa terbaik ditemukan pada Ulasan Aplikasi Gojek yaitu BERT untuk Aspek *Service* akurasi 98,78%, presisi 96%, *recall* 94,73%, dan *F1-Score* 95,36.

Referensi

- [1] Syaifulloh Amien Pandega Perdana, Teguh Bharata Aji, and Ridi Ferdiana, "Aspect Category Classification dengan Pendekatan Machine Learning Menggunakan Dataset Bahasa Indonesia," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 10, no. 3, 2021, doi: 10.22146/jnteti.v10i3.1819.
- [2] J. W. Iskandar and Y. Nataliani, "Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 6, 2021, doi: 10.29207/resti.v5i6.3588.
- [3] H. Mustakim and S. Priyanta, "Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 16, no. 2, 2022, doi: 10.22146/ijccs.68903.
- [4] Salsabila, S. M. P. Tyas, Y. Romadhona, and D. Purwitasari, "Aspect-based Sentiment and Correlation-based Emotion Detection on Tweets for Understanding Public Opinion of Covid-19.," *Journal of Information Systems Engineering and Business Intelligence*, vol. 9, no. 1, 2023, doi: 10.20473/jisebi.9.1.84-94.
- [5] C. A. Bahri and L. H. Suadaa, "Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. 1, 2023, doi: 10.22146/ijccs.77354.
- [6] D. D. A. Yani, H. S. Pratiwi, and H. Muhandi, "Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace," *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*, vol. 7, no. 4, 2019, doi: 10.26418/justin.v7i4.30930.
- [7] S. A. Azzahra and A. Wibowo, "Analisis Sentimen Multi-Aspek Berbasis Konversi Ikon Emosi dengan Algoritme Naïve Bayes untuk Ulasan Wisata Kuliner Pada Web Tripadvisor," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 4, p. 737, Aug. 2020, doi: 10.25126/jtiik.2020731907.
- [8] A. Hermawan, I. Jowensen, J. Junaedi, and Edy, "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine," *JST (Jurnal Sains dan Teknologi)*, vol. 12, no. 1, pp. 129–137, Apr. 2023, doi: 10.23887/jstundiksha.v12i1.52358.
- [9] Y. Alotaibi *et al.*, "Suggestion mining from opinionated text of big social media data," *Computers, Materials and Continua*, vol. 68, no. 3, 2021, doi: 10.32604/cmc.2021.016727.

- [10] N. Z. Dina and N. Juniarta, "Aspect based Sentiment Analysis of Employee's Review Experience," *Journal of Information Systems Engineering and Business Intelligence*, vol. 6, no. 1, 2020, doi: 10.20473/jisebi.6.1.79-88.
- [11] M. Qorib, T. Oladunni, M. Denis, E. Ososanya, and P. Cotae, "Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset," *Expert Syst Appl*, vol. 212, 2023, doi: 10.1016/j.eswa.2022.118715.
- [12] M. P. Geetha and D. Karthika Renuka, "Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model," *International Journal of Intelligent Networks*, vol. 2, pp. 64–69, 2021, doi: 10.1016/j.ijin.2021.06.005.
- [13] F. Rashif, G. Ihza Perwira Nirvana, M. Alif Noor, and N. Aini Rakhmawati, "Implementasi LDA untuk Pengelompokan Topik Cuitan Akun Bot Twitter bertagar #Covid-19 LDA Implementation for Topic of Bot's Tweets with #Covid-19 Hashtag," *Cogito Smart Journal /*, vol. 7, no. 1, 2021.
- [14] W. A. Nawang Sari and H. Dwi Purnomo, "TOPIC MODELING USING THE LATENT DIRICHLET ALLOCATION METHOD ON WIKIPEDIA PANDEMIC COVID-19 DATA IN INDONESIA," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 5, 2022, doi: 10.20884/1.jutif.2022.3.5.321.
- [15] Y. Sahria and D. H. Fudholi, "Analisis Topik Penelitian Kesehatan di Indonesia Menggunakan Metode Topic Modeling LDA," *Jurnal Rekayasa Sistem dan Teknologi Informasi*, vol. 4, no. 2, 2020.
- [16] R. Srivastava, P. K. Bharti, and P. Verma, "Comparative Analysis of Lexicon and Machine Learning Approach for Sentiment Analysis," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 3, 2022, doi: 10.14569/IJACSA.2022.0130312.
- [17] J. He, L. Xu, Y. Fan, Z. Xu, M. Yan, and Y. Lei, "Deep Learning Based Valid Bug Reports Determination and Explanation," in *2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE)*, IEEE, Oct. 2020, pp. 184–194. doi: 10.1109/ISSRE5003.2020.00026.