

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI FLOQ  
SEBAGAI PLATFORM INVESTASI KRIPTO DI INDONESIA  
MENGUNAKAN INDOBERT****Bintang Aqil Mahendra<sup>1\*</sup>, Agung Panji Sasmito<sup>2</sup>, Deddy Rudhistiar<sup>3</sup>**<sup>1,2</sup>Program Studi Bisnis Digital, Institut Teknologi Nasional Malang<sup>3</sup>Program Studi Teknik Informatika, Institut Teknologi Nasional Malangemail: bintangaqilm@gmail.com<sup>1\*</sup>

**Abstrak:** Meningkatnya adopsi platform investasi aset kripto di Indonesia memerlukan evaluasi sistematis terhadap respons pengguna melalui pendekatan analisis sentimen. Penelitian ini menganalisis opini publik terhadap aplikasi FLOQ sebagai platform investasi kripto lokal dengan memanfaatkan model IndoBERT yang telah di-*fine-tune*. Dataset penelitian terdiri dari 6.669 ulasan berbahasa Indonesia yang diperoleh dari Google Play Store pada periode Mei hingga Oktober 2025 menggunakan metode *web scraping*. Pelabelan awal sentimen dilakukan dengan model RoBERTa W11WO, kemudian hasil pelabelan tersebut digunakan dalam proses *fine-tuning* model IndoBERT dengan tiga skema pembagian data, yaitu 70:30, 80:20, dan 90:10. Evaluasi kinerja model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa skema pembagian data 70:30 menghasilkan performa terbaik dengan nilai *accuracy* sebesar 93,95% dan *F1-score* sebesar 0,87. Distribusi sentimen menunjukkan dominasi opini positif sebesar 76,1%, diikuti oleh sentimen negatif sebesar 13,4% dan netral sebesar 10,5%. Analisis aspek bisnis terhadap lima dimensi layanan menunjukkan bahwa aspek *usability* memperoleh sentimen positif tertinggi sebesar 92,63%, sementara aspek Keamanan memiliki proporsi sentimen negatif tertinggi sebesar 18,15%. Temuan ini menunjukkan bahwa model IndoBERT efektif dalam mengklasifikasikan sentimen ulasan berbahasa Indonesia serta memberikan masukan berbasis data bagi pengembangan layanan dan fitur aplikasi FLOQ sebagai platform investasi aset kripto lokal.

**Kata Kunci :** analisis sentimen; *fine-tuning*; FLOQ; IndoBERT; investasi kripto

**PENDAHULUAN**

Perkembangan platform investasi aset kripto di Indonesia mendorong munculnya berbagai aplikasi lokal yang berupaya memberikan kemudahan akses bagi masyarakat, khususnya investor pemula. Aplikasi FLOQ dikembangkan sebagai platform investasi kripto lokal yang tidak hanya menyediakan layanan transaksi aset digital, tetapi juga mengedepankan pendekatan edukatif dan kemudahan penggunaan bagi penggunanya. Keberadaan fitur edukasi dan antarmuka yang sederhana menjadi strategi penting dalam meningkatkan kepercayaan pengguna terhadap aplikasi investasi kripto di tengah tingginya risiko dan volatilitas pasar aset digital [1], [2].

Ulasan pengguna pada aplikasi digital mencerminkan pengalaman nyata serta persepsi pengguna terhadap kualitas layanan yang diberikan. Pada aplikasi keuangan dan investasi, ulasan pengguna sering memuat opini mengenai kemudahan penggunaan, proses layanan, dan keandalan sistem, yang menjadi indikator penting dalam mengevaluasi kinerja aplikasi. Namun, jumlah ulasan yang terus bertambah membuat proses analisis secara manual menjadi tidak efisien. Oleh karena itu, pendekatan *sentiment analysis* berbasis *natural language processing* banyak digunakan untuk mengelompokkan opini pengguna secara otomatis ke dalam kategori sentimen seperti positif, negatif, dan netral. Survei terkini menunjukkan bahwa *sentiment analysis* telah diterapkan secara luas untuk menganalisis data opini dalam skala besar dan membantu memahami kecenderungan persepsi pengguna terhadap suatu sistem atau layanan [3]. Penerapan pendekatan ini pada ulasan aplikasi di Indonesia juga terbukti mampu memberikan gambaran pola sentimen pengguna yang relevan sebagai dasar evaluasi dan pengembangan aplikasi digital [4].

Namun, bertambahnya jumlah ulasan pengguna menyebabkan analisis manual menjadi tidak efisien serta berisiko menimbulkan subjektivitas. Untuk mengatasi permasalahan tersebut, *sentiment analysis* berbasis *Natural Language Processing (NLP)* banyak diterapkan guna mengklasifikasikan opini pengguna secara otomatis ke dalam kategori sentimen tertentu. Pendekatan ini dinilai mampu mengolah data teks dalam skala besar secara konsisten dan objektif, sehingga efektif digunakan untuk menganalisis persepsi pengguna terhadap aplikasi digital dan layanan berbasis teknologi [5], [6].

Model *Bidirectional Encoder Representations from Transformers (BERT)* merupakan pendekatan berbasis arsitektur *Transformer* yang dikembangkan untuk memahami konteks teks secara dua arah dan terbukti efektif dalam berbagai tugas *Natural Language Processing*, termasuk klasifikasi sentimen [7]. Arsitektur *Transformer* yang menjadi dasar pengembangan BERT memungkinkan model menangkap hubungan kontekstual antar kata secara lebih optimal dibandingkan pendekatan sekuensial konvensional [8]. Untuk konteks bahasa Indonesia, IndoBERT dikembangkan menggunakan korpus berbahasa Indonesia dalam skala besar sehingga mampu merepresentasikan karakteristik linguistik lokal secara lebih akurat dibandingkan model multibahasa [9].

Selain klasifikasi sentimen secara umum, analisis terhadap aspek layanan tertentu juga diperlukan untuk memperoleh pemahaman yang lebih mendalam. Analisis berbasis aspek memungkinkan identifikasi bagian layanan yang menjadi kekuatan maupun kelemahan suatu aplikasi berdasarkan opini pengguna [10]. Dalam penelitian ini, analisis aspek



bisnis dilakukan menggunakan pendekatan *rule-based classification* dengan *keyword matching*, sehingga setiap ulasan dapat dikaitkan dengan aspek layanan tertentu tanpa menerapkan metode *aspect extraction* otomatis berbasis pembelajaran mesin.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi FLOQ sebagai platform investasi aset kripto lokal menggunakan model IndoBERT yang telah di-*fine-tune*. Selain itu, penelitian ini juga menyajikan analisis aspek bisnis untuk mengidentifikasi dimensi layanan yang memperoleh respons positif maupun negatif dari pengguna. Analisis sentimen dan aspek tersebut digunakan untuk memetakan persepsi pengguna terhadap layanan aplikasi FLOQ secara komprehensif berdasarkan ulasan yang diberikan. Model IndoBERT dimanfaatkan sebagai metode klasifikasi untuk memperoleh hasil analisis yang optimal, sementara evaluasi performa model dilakukan sebagai pendukung untuk memastikan keandalan hasil. Dengan demikian, kontribusi utama penelitian ini terletak pada penyajian analisis persepsi pengguna FLOQ berbasis data ulasan, yang diharapkan dapat menjadi masukan bagi pengembang dalam meningkatkan kualitas layanan dan fitur aplikasi.

## TINJAUAN PUSTAKA

Analisis sentimen adalah cabang dari *text mining* yang bertujuan mengidentifikasi orientasi opini pada data teks, seperti komentar atau ulasan pengguna, menjadi kategori polaritas sentimen (positif, negatif, netral) berdasarkan pola linguistik dalam kalimat [11]. Pendekatan ini penting dalam evaluasi persepsi pengguna terhadap layanan aplikasi digital karena mampu memetakan opini massal secara otomatis dengan akurasi yang lebih tinggi dibandingkan metode manual.

Dalam *NLP*, teks diproses melalui tahapan seperti *tokenization*, pembersihan teks (*cleaning*), penghapusan *stopwords*, dan *normalization* untuk menghasilkan representasi teks yang siap dimodelkan [12]. Perkembangan *deep learning*, terutama arsitektur *Transformer*, memperkenalkan mekanisme *self-attention* yang mampu menangkap hubungan konteks antar kata dalam satu kalimat atau dokumen tanpa urutan sekuensial, sehingga memperbaiki kemampuan model memahami makna bahasa secara keseluruhan [8].

Berdasarkan arsitektur tersebut, model *Bidirectional Encoder Representations from Transformers (BERT)* dilatih untuk memahami konteks dua arah kalimat sehingga representasi fitur bahasa yang dihasilkan lebih kaya daripada model sebelumnya. *IndoBERT* merupakan adaptasi BERT yang dilatih khusus pada korpus berbahasa Indonesia sehingga lebih unggul dalam menangani teks ulasan Indonesia [13]. ntuk tugas spesifik seperti klasifikasi sentimen, model pra-latih ini diadaptasi melalui *fine-tuning*, yaitu pelatihan lanjutan menggunakan data berlabel yang sesuai dengan domain aplikasi. Pendekatan *fine-tuning* terbukti mampu meningkatkan kinerja klasifikasi teks berbahasa Indonesia karena memungkinkan model menyesuaikan parameter dengan karakteristik linguistik dan konteks ulasan pengguna secara lebih spesifik [14].

Selain klasifikasi sentimen global, analisis sentimen berbasis aspek memungkinkan pemetaan opini pada dimensi fitur tertentu dari layanan, seperti pengalaman penggunaan atau keamanan. Pendekatan aspek ini memberikan wawasan yang lebih terperinci terhadap bagian layanan yang menjadi perhatian pengguna [15]. Dalam penelitian ini, pemetaan aspek bisnis dilakukan menggunakan *rule-based classification* dengan *keyword matching*, yaitu klasifikasi berdasarkan kemunculan kata kunci terkait aspek yang telah ditentukan sebelumnya, efektif untuk analisis aspek yang telah terdefinisi meskipun sensitivitasnya bergantung pada kelengkapan daftar kata kunci [16].

Beberapa penelitian terdahulu menunjukkan bahwa *IndoBERT* efektif digunakan untuk analisis sentimen pada berbagai konteks di Indonesia. Analisis sentimen publik pada media sosial X terkait Badan Investasi Danantara menunjukkan bahwa *IndoBERT* mampu mencapai akurasi sebesar 88,77% dan *F1-score* sebesar 88,76%, dengan mayoritas sentimen negatif sebesar 58,6% [17]. Penelitian lain pada ulasan pengguna aplikasi *telemedicine* menunjukkan bahwa penerapan *fine-tuning IndoBERT* dengan rasio pembagian data 90:10 menghasilkan performa terbaik dengan akurasi 83,71%, lebih tinggi dibandingkan rasio 80:20 dan 70:30 [18].

Selain itu, penggunaan *IndoBERT* pada analisis sentimen media sosial Twitter terhadap kinerja Penjabat Gubernur DKI Jakarta menghasilkan performa sangat baik dengan akurasi, *precision*, *recall*, dan *F1-score* masing-masing mencapai 96% [19]. Pada konteks berbeda, analisis sentimen komentar YouTube terhadap program Tabungan Perumahan Rakyat (TAPERA) menggunakan *IndoBERT Lite-Large* menghasilkan akurasi sebesar 78% dari total 9.571 data, dengan dominasi sentimen negatif [20]. Meskipun demikian, penelitian-penelitian tersebut masih berfokus pada klasifikasi sentimen secara umum dan belum mengaitkannya secara eksplisit dengan analisis aspek layanan, sehingga membuka peluang penelitian lanjutan pada domain aplikasi investasi kripto lokal seperti FLOQ.

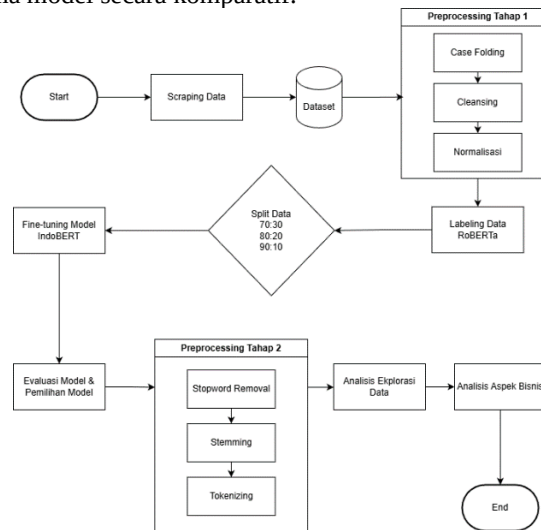
Dengan demikian, gap yang masih ada adalah kurangnya kajian yang memadukan analisis sentimen berbasis *IndoBERT* dan pemetaan aspek bisnis menggunakan pendekatan aturan pada ulasan aplikasi investasi kripto lokal seperti FLOQ. Penelitian ini bertujuan mengisi celah tersebut dengan menggabungkan kedua pendekatan untuk memperoleh wawasan yang lebih komprehensif terhadap persepsi pengguna.

## METODE

Penelitian ini mengadopsi pendekatan eksperimental untuk menganalisis kinerja model analisis sentimen IndoBERT dalam mengklasifikasikan opini pengguna terhadap aplikasi FLOQ. Fokus utama penelitian adalah pada klasifikasi sentimen ulasan ke dalam tiga kategori, yaitu positif, negatif, dan netral. Metodologi penelitian ini dirancang untuk mengevaluasi secara sistematis pengaruh perbedaan rasio pembagian data latih dan data uji terhadap kemampuan



model IndoBERT dalam mengidentifikasi sentimen pada teks ulasan berbahasa Indonesia yang bersifat informal dan bervariasi. Tahapan penelitian disusun mengikuti alur visual pada Gambar 1, dengan rangkaian proses yang meliputi pengumpulan data, pra-pemrosesan data, pelabelan data secara otomatis, pembagian dataset, *fine-tuning* model IndoBERT, serta evaluasi performa model secara komparatif.



Gambar 1. Flowchart Tahapan Penelitian

Gambar 1 menunjukkan alur metodologi penelitian yang digunakan dalam penelitian ini. Proses penelitian dimulai dari pengumpulan data ulasan pengguna, dilanjutkan dengan tahap *preprocessing*, pelabelan data, pembagian dataset, pelatihan dan evaluasi model IndoBERT, hingga analisis eksplorasi data dan analisis aspek bisnis. Setiap tahapan pada alur tersebut saling berkaitan dan dirancang untuk menghasilkan analisis sentimen yang komprehensif terhadap aplikasi FLOQ.

### Pengumpulan Data

Data ulasan aplikasi FLOQ dikumpulkan dari Google Play Store menggunakan teknik *web scraping* dengan *library google-play-scraper* berbasis Python. Pengambilan data difokuskan pada ulasan berbahasa Indonesia dalam rentang waktu 30 Mei hingga 31 Oktober 2025, yang merepresentasikan fase awal peluncuran dan pertumbuhan aplikasi FLOQ. Proses pengambilan data dilakukan secara bertahap untuk menghindari pembatasan akses dari server Google Play Store. Dari proses ini diperoleh sebanyak 6.672 ulasan yang disimpan dalam format CSV untuk dianalisis lebih lanjut.

### Preprocessing Tahap 1

Tahap *preprocessing* pertama dilakukan untuk membersihkan dan menyeragamkan teks ulasan sebelum proses pelabelan data. Pada tahap ini diterapkan *case folding* untuk mengubah seluruh teks menjadi huruf kecil, *cleansing* untuk menghapus elemen yang tidak relevan seperti URL, tanda baca, emoji, dan angka, serta normalisasi kata untuk mengubah kata tidak baku menjadi bentuk baku menggunakan kamus slang bahasa Indonesia dan kamus kustom. Tahapan ini bertujuan mengurangi *noise* pada data sekaligus mempertahankan struktur kalimat agar informasi kontekstual tetap terjaga.

Penelitian ini tidak menerapkan *stemming* dan *tokenization* pada tahap *preprocessing* pertama, sejalan dengan temuan [21] yang menunjukkan bahwa model IndoBERT dapat memberikan hasil yang lebih optimal tanpa tahapan tersebut. Oleh karena itu, pada tahap awal hanya digunakan *case folding*, *cleansing*, dan normalisasi untuk menjaga konteks kalimat tetap utuh sebelum proses pelabelan data dan *fine-tuning* model.

### Pelabelan Data

Pelabelan data dilakukan secara otomatis menggunakan model *w11wo/indonesian-roberta-base-sentiment-classifier* untuk mengklasifikasikan setiap ulasan ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Model menghasilkan probabilitas untuk masing-masing kelas, kemudian label dengan nilai tertinggi dipilih sebagai hasil pelabelan. Pendekatan pelabelan otomatis ini digunakan sebagai *pseudo-label* untuk mendukung proses pelatihan dan *fine-tuning* model IndoBERT, sehingga memungkinkan pengolahan dataset berukuran besar secara efisien tanpa pelabelan manual.

### Pembagian Data

Dataset yang telah melalui proses pelabelan kemudian dibagi menjadi data latih dan data uji menggunakan beberapa skema rasio pembagian, yaitu 70:30, 80:20, dan 90:10. Pembagian data dilakukan secara *stratified* untuk menjaga proporsi kelas sentimen pada masing-masing subset. Variasi rasio pembagian ini digunakan untuk mengevaluasi



konsistensi kinerja model pada proporsi data pelatihan yang berbeda, sedangkan rincian jumlah data pada setiap skema disajikan pada bagian hasil dan pembahasan.

### Fine-Tuning Model IndoBERT

Proses *fine-tuning* dilakukan menggunakan model *indobenchmark/indobert-base-p1* sebagai model pra-latih bahasa Indonesia. Konfigurasi *hyperparameter* yang digunakan mengacu pada pengaturan fine-tuning model BERT sebagaimana dijelaskan oleh Devlin et al., yang merekomendasikan *learning rate* kecil dan jumlah *epoch* terbatas untuk menjaga stabilitas pelatihan model transformer [15]. Parameter yang digunakan meliputi *learning rate* sebesar  $2e-5$ , *batch size* 16, jumlah *epoch* sebanyak 3, serta *maximum sequence length* 512. Pemilihan jumlah *epoch* yang relatif kecil dilakukan untuk menghindari *overfitting*, mengingat model pra-latih telah memiliki representasi bahasa yang kuat. Model yang telah melalui proses fine-tuning selanjutnya dievaluasi untuk menilai kemampuan klasifikasi sentimen pada data uji.

### Evaluasi Model

Tabel 1. Confusion Matrix Klasifikasi Sentimen

| Kelas<br>Aktual | Kelas Prediksi         |                        |                        |
|-----------------|------------------------|------------------------|------------------------|
|                 | Positif                | Negatif                | Netral                 |
| Positif         | True Positive<br>(TP)  | False Negative<br>(FN) | False Negative<br>(FN) |
| Negatif         | False Positive<br>(FP) | True Negative<br>(TN)  | False Positive<br>(FP) |
| Netral          | False Positive<br>(FP) | False Negative<br>(FN) | True Netral<br>(TN)    |

Evaluasi model dilakukan menggunakan *confusion matrix* untuk membandingkan hasil prediksi model IndoBERT dengan label aktual pada data uji. Pendekatan ini digunakan untuk melihat distribusi prediksi benar dan salah pada masing-masing kelas sentimen, yaitu positif, negatif, dan netral. Informasi yang diperoleh dari *confusion matrix* kemudian digunakan sebagai dasar perhitungan metrik evaluasi yang umum digunakan dalam penelitian klasifikasi sentimen teks, meliputi *accuracy*, *precision*, *recall*, dan *F1-score* [12], [17]. Penggunaan metrik tersebut bertujuan untuk memberikan gambaran performa model secara menyeluruh serta menilai keseimbangan kemampuan model dalam mengklasifikasikan setiap kelas sentimen.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Sebagai pembandingan performa, penelitian ini juga menguji model *baseline* sederhana menggunakan algoritma Naive Bayes dengan representasi fitur TF-IDF (*Term Frequency-Inverse Document Frequency*). Model *baseline* dilatih dan diuji menggunakan rasio pembagian data terbaik, sama seperti konfigurasi terbaik pada model IndoBERT, guna memastikan perbandingan yang adil. Pemilihan Naive Bayes sebagai *baseline* didasarkan pada penggunaan algoritma ini yang umum dalam penelitian klasifikasi sentimen teks, sehingga dapat menjadi acuan untuk mengukur peningkatan performa model berbasis Transformer.

### Inferensi Model

Setelah model IndoBERT dievaluasi dan menunjukkan performa yang stabil, tahap selanjutnya adalah melakukan inferensi terhadap seluruh data ulasan aplikasi FLOQ. Pada tahap ini, model yang telah melalui proses *fine-tuning* digunakan untuk memprediksi sentimen setiap ulasan secara menyeluruh. Hasil inferensi ini menghasilkan label sentimen akhir yang digunakan sebagai dasar dalam analisis lanjutan, sehingga seluruh dataset merefleksikan keputusan model yang telah disesuaikan dengan karakteristik data penelitian.

### Preprocessing Tahap 2

*Preprocessing* tahap kedua dilakukan setelah proses inferensi model dan bertujuan untuk mendukung analisis eksplorasi data serta analisis aspek bisnis. Berbeda dengan *preprocessing* tahap pertama yang difokuskan pada menjaga konteks kalimat untuk keperluan pelabelan dan *fine-tuning*, tahap ini diarahkan pada ekstraksi informasi tekstual yang lebih spesifik. Pada tahap kedua diterapkan *tokenization*, *stopword removal*, dan *stemming* untuk menyederhanakan bentuk kata dan mengurangi kata-kata yang tidak memiliki makna analitis. Proses ini bertujuan untuk menonjolkan kata kunci yang merepresentasikan topik dan aspek yang paling sering dibahas oleh pengguna, sehingga memudahkan visualisasi seperti *wordcloud* serta proses pencocokan kata kunci pada analisis aspek bisnis.



### Analisis Eksplorasi Data

Analisis eksplorasi data dilakukan untuk memperoleh gambaran umum mengenai distribusi dan karakteristik sentimen ulasan aplikasi FLOQ berdasarkan hasil inferensi model IndoBERT. Analisis ini mencakup distribusi jumlah sentimen positif, negatif, dan netral, serta identifikasi kata-kata yang sering muncul pada masing-masing kategori sentimen. Tahap ini bertujuan untuk membantu memahami pola umum respons pengguna sebelum dilakukan analisis yang lebih spesifik terhadap aspek bisnis aplikasi.

### Analisis Aspek Bisnis

Tahap akhir dalam penelitian ini adalah analisis aspek bisnis untuk mengidentifikasi bagian layanan aplikasi FLOQ yang paling sering menjadi perhatian pengguna. Analisis aspek dilakukan menggunakan pendekatan *aspect category classification* berbasis aturan (*rule-based classification*) dengan teknik *keyword matching*. Pendekatan ini dipilih karena mampu memetakan opini pengguna ke dalam kategori aspek tertentu secara terstruktur tanpa memerlukan data latih tambahan, serta efektif digunakan pada ulasan berbahasa Indonesia dengan jumlah data besar [16].

Penentuan aspek bisnis dalam penelitian ini didasarkan pada kajian literatur penelitian terdahulu pada domain aplikasi keuangan digital. Berdasarkan temuan tersebut, aspek yang paling sering digunakan dan relevan meliputi keamanan, *usability*, fitur aplikasi, layanan pelanggan, serta transaksi dan pembayaran [2], [10], [22]. Pemilihan lima aspek ini bertujuan untuk merepresentasikan dimensi utama pengalaman pengguna pada aplikasi investasi digital serta memungkinkan analisis sentimen yang lebih terarah dibandingkan analisis sentimen secara umum. Dalam proses analisis, satu ulasan pengguna dimungkinkan untuk diklasifikasikan ke dalam lebih dari satu aspek apabila ulasan tersebut membahas beberapa dimensi layanan secara bersamaan, sebagaimana dijelaskan dalam penelitian *aspect-based sentiment analysis* yang menyatakan bahwa satu ulasan dapat mengandung lebih dari satu kategori aspek [23].

## HASIL DAN PEMBAHASAN

### Pengumpulan Data

Data ulasan aplikasi FLOQ dikumpulkan dari Google Play Store pada periode Mei hingga Oktober 2025. Periode ini mencerminkan fase awal penggunaan aplikasi oleh masyarakat, sehingga mampu menggambarkan persepsi pengguna sejak tahap awal adopsi hingga penggunaan berkelanjutan. Dari proses pengumpulan data, diperoleh 6.672 ulasan pengguna yang selanjutnya digunakan sebagai data mentah penelitian.

### Preprocessing Data Tahap 1

Pada tahap preprocessing pertama, data ulasan yang telah dikumpulkan mengalami proses pembersihan dan penyeragaman. Dari total 6.672 ulasan, sebanyak 6.669 ulasan dinyatakan layak digunakan, sementara sisanya dieliminasi karena duplikasi, teks kosong, atau tidak dapat diproses secara optimal. Hasil ini menunjukkan bahwa sebagian besar data yang dikumpulkan memiliki kualitas yang baik untuk dianalisis lebih lanjut.

Tabel 2. Preprocessing Tahap 1

| Tahap         | Hasil                                                                                                                                                                                                                                                        |
|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Original Teks | #floqarjuna Aplikasi Keren Menurut gua, simple aksesnya, gak bnyak konten konten yg buat bingung pengunjung baru, aplikasinya juga enteng banget, buat hp hp jadul kyak 2020 aja masih sat set sat set klik nya, pokok nya super enteng dan gacor dahhh 🤔🤔🤔  |
| Case Folding  | #floqarjuna aplikasi keren menurut gua, simple aksesnya, gak banyak konten konten yg buat bingung pengunjung baru, aplikasinya juga enteng banget, buat hp hp jadul kyak 2020 aja masih sat set sat set klik nya, pokok nya super enteng dan gacor dahhh 🤔🤔🤔 |
| Cleansing     | aplikasi keren menurut gua simple aksesnya gak bnyak konten konten yg buat bingung pengunjung baru aplikasinya juga enteng banget buat hp hp jadul kyak 2020 aja masih sat set sat set klik nya pokok nya super enteng dan gacor dahhh                       |
| Normalisasi   | aplikasi keren menurut gua sederhana aksesnya tidak banyak konten konten yang buat bingung pengunjung baru aplikasinya juga enteng banget buat hp hp jadul kayak 2020 saja masih sat set sat set klik nya pokok nya super enteng dan gacor dahhh             |

### Pelabelan Data Sentimen

Hasil pelabelan sentimen terhadap ulasan pengguna aplikasi FLOQ menunjukkan bahwa dari total 6.669 ulasan yang dianalisis, sebanyak 5.034 ulasan (75,48%) diklasifikasikan sebagai sentimen positif. Selanjutnya, sebanyak 901 ulasan (13,51%) termasuk dalam sentimen negatif, sedangkan 734 ulasan (11,01%) dikategorikan sebagai sentimen netral.

Tabel 3. Labeling Data

| Teks                                                                                                                                      | Label   |
|-------------------------------------------------------------------------------------------------------------------------------------------|---------|
| aplikasi jual beli kripto yang terpercaya dan cepat proses registrasi mudah tampilannya modern dan aman banget buat transaksi rekomendasi | Positif |
| tolong tambahkan untuk setoran tambahkan e walet dana kalau sudah di tambahkan e walet dana gue janji bakal pembaruan rating bintang      | Netral  |





floq keren banget buat perdagangan kripto tampilannya simpel dan mudah dipakai transaksi juga cepat tanpa rumit wajib coba deh  
bisa enggak kalau verifikasi pakai sim ktp belum pembaruan malas mengurus masih banyak oknum

Positif

Negatif

Untuk memastikan keandalan hasil pelabelan sentimen otomatis tersebut, dilakukan validasi manual terhadap sebagian data ulasan sebagai *ground truth*. Sebanyak 300 ulasan dipilih secara acak dari total 6.669 data dan diberi label sentimen secara manual ke dalam tiga kelas, yaitu positif, netral, dan negatif. Hasil anotasi manual kemudian dibandingkan dengan hasil pelabelan sentimen otomatis yang diperoleh menggunakan model RoBERTa. Berdasarkan hasil evaluasi, diperoleh tingkat akurasi sebesar 87,33%, yang menunjukkan bahwa pelabelan sentimen otomatis yang digunakan cukup andal dan layak digunakan sebagai dasar dalam proses analisis sentimen pada penelitian ini.

### Hasil Fine-Tuning dan Evaluasi Model

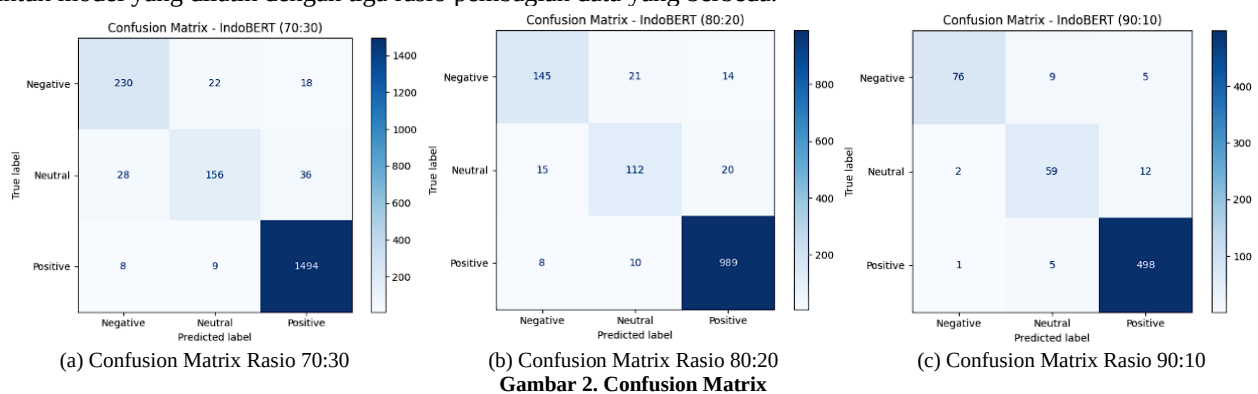
Hasil evaluasi model menunjukkan bahwa model IndoBERT mampu mengklasifikasikan sentimen ulasan dengan performa yang konsisten pada seluruh skema pembagian data. Nilai accuracy pada masing-masing skema berada pada rentang yang relatif berdekatan, sehingga menunjukkan stabilitas model dalam mengenali pola sentimen pada data ulasan aplikasi FLOQ.

Tabel 4. Evaluasi Model IndoBERT

| Split Ratio | Train Size | Test Size | Accuracy (%) | Precision (Avg) | Recall (Avg) | F1-Score (Avg) |
|-------------|------------|-----------|--------------|-----------------|--------------|----------------|
| 70:30       | 4668       | 2.001     | 93,95        | 0,89            | 0,85         | 0,87           |
| 80:20       | 5535       | 1.334     | 93,40        | 0,87            | 0,85         | 0,86           |
| 90:10       | 6.002      | 667       | 94.90        | 0,91            | 0,88         | 0,90           |

Hasil pada Tabel 4 menunjukkan bahwa terdapat perbedaan karakteristik performa model pada setiap skema pembagian data yang diuji. Model yang dilatih dengan rasio 90:10 mencapai metrik evaluasi tertinggi dengan nilai F1-score sebesar 0,90, diikuti oleh model dengan rasio 70:30 yang mencatat F1-score 0,87, dan rasio 80:20 dengan F1-score 0,86. Meskipun rasio 90:10 menghasilkan nilai akurasi tertinggi sebesar 94,90%, performa ini perlu dianalisis lebih lanjut dengan mempertimbangkan ukuran data uji yang relatif kecil yaitu hanya 667 sampel. Ukuran data uji yang terbatas dapat menyebabkan evaluasi performa menjadi kurang stabil dan cenderung optimistik karena tidak merepresentasikan variasi data secara menyeluruh.

Untuk analisis yang lebih mendalam, confusion matrix digunakan untuk mengevaluasi kemampuan model dalam membedakan antar kelas sentimen pada masing-masing konfigurasi split data. Gambar 2 menyajikan confusion matrix untuk model yang dilatih dengan tiga rasio pembagian data yang berbeda.



Confusion matrix pada Gambar 2 menunjukkan pola klasifikasi yang konsisten di mana model memiliki performa terbaik pada kelas positif yang merupakan kelas mayoritas dalam dataset. Model dengan *split* 70:30 berhasil mengklasifikasikan 1.494 data positif dengan benar, namun masih terdapat kesalahan pada kelas netral di mana 36 data netral salah diprediksi sebagai positif. Kesalahan klasifikasi terbesar terjadi pada transisi antara kelas netral dan positif, yang mengindikasikan bahwa model kesulitan membedakan ulasan bersifat informatif netral dengan ulasan yang mengandung apresiasi implisit karena karakteristik sentimen netral yang cenderung ambigu dan memiliki kedekatan semantik dengan frasa deskriptif positif.

Berdasarkan hasil eksperimen pada tiga konfigurasi pembagian data, model dengan *split* 70:30 menunjukkan performa klasifikasi yang paling seimbang. Meskipun *split* 90:10 menghasilkan nilai *accuracy* tertinggi sebesar 94,90%, *split* 70:30 memberikan kombinasi metrik yang lebih stabil dengan nilai *accuracy* sebesar 93,95%, *precision* sebesar 0,89, *recall* sebesar 0,85, dan *F1-score* sebesar 0,87. Keseimbangan nilai *precision*, *recall*, dan *F1-score* pada *split* 70:30 menunjukkan bahwa model mampu mengklasifikasikan setiap kelas sentimen secara lebih proporsional. Temuan tersebut menunjukkan bahwa pemilihan rasio pembagian data tidak hanya memengaruhi nilai akurasi, tetapi juga berdampak pada



stabilitas dan reliabilitas evaluasi performa model secara keseluruhan. Hasil tersebut sejalan dengan temuan [24] yang menunjukkan bahwa pengujian beberapa rasio pembagian data menghasilkan performa terbaik pada *split* 70:30, khususnya pada metrik *F1-score*, dibandingkan rasio lainnya. Penelitian [25] juga menjelaskan bahwa penggunaan rasio dengan porsi data uji yang terlalu kecil seperti pada *split* 90:10 dapat menyebabkan evaluasi performa model menjadi kurang stabil dan cenderung optimistik karena ukuran data uji yang terbatas kurang merepresentasikan variasi data secara keseluruhan sehingga berpotensi meningkatkan risiko *overfitting* pada tahap evaluasi model. Dengan demikian, pemilihan *split* 70:30 sebagai model terbaik dalam penelitian ini didasarkan pada pertimbangan performa klasifikasi yang lebih seimbang serta reliabilitas evaluasi yang lebih baik, sehingga model IndoBERT dengan *split* 70:30 dinilai paling representatif dan layak digunakan sebagai model final dalam sistem klasifikasi sentimen ulasan aplikasi FLOQ.

Sebagai validasi tambahan terhadap performa model IndoBERT, dilakukan perbandingan dengan model baseline Naive Bayes yang diuji pada konfigurasi data yang sama (*split* 70:30). Hasil perbandingan disajikan pada Tabel 5.

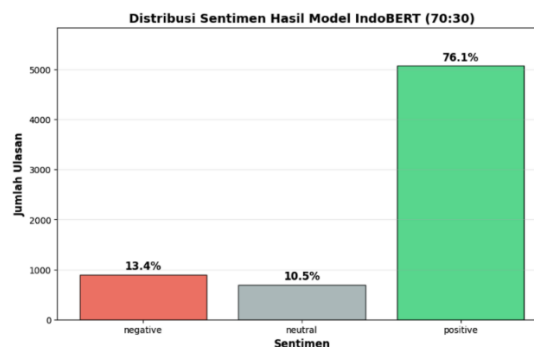
Tabel 5. Perbandingan Performa IndoBERT dan Naive Bayes

| Model               | Accuracy (%) | Precision | Recall | F1-Score |
|---------------------|--------------|-----------|--------|----------|
| IndoBERT (70:30)    | 93,95        | 0,89      | 0,85   | 0,87     |
| Naive Bayes (70:30) | 86,66        | 0,84      | 0,63   | 0,84     |

Berdasarkan Tabel 5, model IndoBERT menunjukkan performa yang lebih tinggi dibandingkan Naive Bayes pada seluruh metrik evaluasi, dengan selisih accuracy sebesar 7,29%. Perbedaan yang paling signifikan terlihat pada nilai recall (0,85 vs 0,63), yang menunjukkan bahwa IndoBERT lebih mampu mengenali kelas minoritas seperti sentimen negatif dan netral. Hal ini membuktikan bahwa pendekatan berbasis Transformer lebih efektif dalam menangkap konteks dan variasi bahasa pada ulasan pengguna dibandingkan metode klasifikasi tradisional berbasis probabilistik.

#### Inferensi Model

Model terbaik selanjutnya digunakan untuk melakukan inferensi terhadap seluruh data ulasan. Dari total 6.669 ulasan yang dianalisis, sebanyak 5.076 ulasan (76.1%) diklasifikasikan sebagai sentimen positif, 896 (13.4%) ulasan sebagai sentimen negatif, dan 697 (10.5%) ulasan sebagai sentimen netral. Distribusi hasil inferensi tersebut ditampilkan pada Gambar 3 dalam bentuk histogram persentase untuk memberikan gambaran visual mengenai sebaran sentimen ulasan pengguna aplikasi FLOQ.



Gambar 3. Distribusi Sentimen Hasil Inferensi Model

#### Preprocessing Data Tahap 2

Preprocessing tahap kedua dilakukan untuk mendukung analisis eksplorasi dan analisis aspek bisnis. Hasil dari tahap ini berupa representasi teks yang lebih ringkas dan fokus pada kata-kata bermakna, sehingga memudahkan identifikasi pola dan topik yang sering muncul dalam ulasan pengguna.

Tabel 6. Preprocessing Tahap 2

| Tahap            | Hasil                                                                                                                                                                                                                                                                                                |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Stopword Removal | aplikasi keren gua sederhana akses banyak konten konten buat bingung unjung baru aplikasi enteng banget buat hp hp jadul kayak masih sat set sat set klik nya pokok nya super enteng gacor dahhh                                                                                                     |
| Stemming         | aplikasi keren gua sederhana akses banyak konten konten buat bingung unjung baru aplikasi enteng banget buat hp hp jadul kayak masih sat set sat set klik nya pokok nya super enteng gacor dahhh                                                                                                     |
| Tokenisasi       | ['aplikasi', 'keren', 'gua', 'sederhana', 'akses', 'banyak', 'konten', 'konten', 'buat', 'bingung', 'unjung', 'baru', 'aplikasi', 'enteng', 'banget', 'buat', 'hp', 'hp', 'jadul', 'kayak', 'masih', 'sat', 'set', 'sat', 'set', 'klik', 'nya', 'pokok', 'nya', 'super', 'enteng', 'gacor', 'dahhh'] |

#### Exploratory Data Analysis

Hasil *exploratory data analysis* menunjukkan perbedaan karakteristik yang jelas antara ulasan positif dan negatif. Ulasan positif didominasi oleh kata-kata yang berkaitan dengan kemudahan penggunaan, kenyamanan, dan keamanan aplikasi. Sebaliknya, ulasan negatif lebih banyak menyoroti kendala teknis, proses verifikasi, dan hambatan transaksi.

(a) Wordcloud Positif

(b) Wordcloud Negatif

Gambar 4 menampilkan visualisasi wordcloud sentimen positif dan negatif pada ulasan pengguna aplikasi FLOQ. Pada wordcloud sentimen positif, kata-kata yang paling menonjol antara lain *investasi*, *saja*, *cepat*, *mudah*, dan *banget*. Kemunculan kata-kata tersebut menunjukkan bahwa ulasan positif banyak membahas kemudahan penggunaan aplikasi, kecepatan akses, serta pengalaman pengguna dalam melakukan aktivitas investasi. Sementara itu, wordcloud sentimen negatif memperlihatkan kata-kata seperti *gagal*, *mau*, *susah*, *lama*, dan *ktp*. Kata-kata tersebut mencerminkan keluhan pengguna yang berkaitan dengan kendala teknis, proses verifikasi akun, serta hambatan dalam penggunaan aplikasi. Perbedaan pola kata yang muncul pada kedua wordcloud tersebut menggambarkan perbedaan fokus pembahasan antara ulasan dengan sentimen positif dan negatif.

**Tren Sentimen per Bulan (IndoBERT)**

| Bulan    | Negative | Neutral | Positive |
|----------|----------|---------|----------|
| May      | ~100     | ~100    | ~300     |
| Jun 2025 | ~500     | ~400    | ~3200    |
| Jul      | ~100     | ~100    | ~500     |
| Aug      | ~50      | ~50     | ~750     |
| Sep      | ~50      | ~50     | ~200     |
| Ok       | ~50      | ~50     | ~100     |

Gambar 5 menampilkan analisis tren sentimen ulasan pengguna aplikasi FLOQ per bulan selama periode pengumpulan data. Terlihat adanya peningkatan jumlah ulasan yang cukup signifikan pada bulan Juni, terutama pada kategori sentimen positif. Peningkatan ini bertepatan dengan periode awal peluncuran aplikasi FLOQ pada akhir Mei, sehingga dapat menggambarkan respons awal pengguna setelah aplikasi tersedia untuk digunakan. Pada bulan Juli, jumlah ulasan pada seluruh kategori sentimen mengalami penurunan, kemudian kembali meningkat pada bulan Agustus meskipun tidak setinggi periode sebelumnya. Selanjutnya, pada bulan September hingga Oktober, jumlah ulasan menunjukkan kecenderungan menurun. Pola ini memberikan gambaran mengenai perubahan aktivitas ulasan pengguna dari waktu ke waktu selama periode penelitian.

## Analisis Aspek Bisnis

Analisis aspek bisnis dilakukan untuk mengidentifikasi bagian layanan aplikasi FLOQ yang paling banyak dibahas oleh pengguna serta melihat kecenderungan sentimen pada setiap aspek. Pada penelitian ini, analisis aspek difokuskan pada lima aspek bisnis, yaitu Keamanan, Kemudahan Penggunaan (*Usability*), Fitur Aplikasi, Layanan Pelanggan (*Customer Service*), serta Transaksi/Pembayaran. Pemetaan ulasan ke dalam aspek dilakukan menggunakan pendekatan *aspect category classification* berbasis *keyword matching*, sehingga hanya ulasan yang secara eksplisit membahas aspek layanan tertentu yang dianalisis. Satu ulasan pengguna dimungkinkan untuk diklasifikasikan ke dalam lebih dari satu aspek apabila ulasan tersebut membahas beberapa layanan sekaligus, sehingga jumlah kemunculan aspek merepresentasikan banyaknya pembahasan aspek, bukan jumlah ulasan secara langsung.

Untuk menilai keandalan hasil pemetaan aspek berbasis aturan, dilakukan evaluasi dengan membandingkan hasil klasifikasi aspek terhadap anotasi manual pada sebagian data ulasan. Sebanyak 300 ulasan dipilih secara acak untuk proses validasi, dengan ketentuan bahwa satu ulasan dapat memiliki lebih dari satu aspek. Hasil klasifikasi dianggap sesuai apabila terdapat minimal satu kesesuaian antara aspek yang dihasilkan oleh metode rule-based dan aspek hasil anotasi manual.

Berdasarkan evaluasi tersebut, dari 300 ulasan yang divalidasi, sebanyak 267 ulasan (89%) berhasil diklasifikasikan dengan benar, sementara 33 ulasan (11%) mengalami ketidaksesuaian. Tingkat akurasi sebesar 89% menunjukkan bahwa pendekatan keyword matching yang digunakan mampu mengidentifikasi aspek yang relevan pada





sebagian besar ulasan pengguna. Ketidakesesuaian klasifikasi umumnya terjadi pada ulasan yang menyampaikan aspek secara implisit atau menggunakan istilah di luar daftar kata kunci yang telah ditetapkan.

Tabel 7. Distribusi Aspek

| Aspek Bisnis         | Total Kemunculan | Persentase |
|----------------------|------------------|------------|
| Usability            | 3.743            | 37,35%     |
| Fitur Aplikasi       | 2.320            | 23,15%     |
| Keamanan             | 2.253            | 22,48%     |
| Transaksi/Pembayaran | 1.352            | 13,49%     |
| Customer Service     | 353              | 3,52%      |

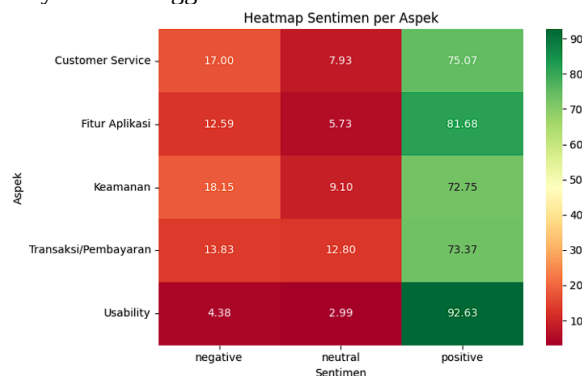
Berdasarkan Tabel 7, aspek Kemudahan Penggunaan merupakan aspek yang paling banyak dibahas oleh pengguna dengan 3.743 ulasan atau sebesar 37,35%. Selanjutnya diikuti oleh aspek Fitur Aplikasi sebanyak 2.320 ulasan (23,15%) dan aspek Keamanan sebanyak 2.253 ulasan (22,48%). Aspek Transaksi dan Pembayaran tercatat sebanyak 1.352 ulasan (13,49%), sedangkan aspek Layanan Pelanggan memiliki jumlah ulasan paling sedikit yaitu 353 ulasan (3,52%). Distribusi ini menunjukkan adanya variasi fokus pembahasan pengguna terhadap aspek-aspek layanan aplikasi FLOQ.

Tabel 8. Distribusi Sentimen per Aspek

| Aspek Bisnis         | Negatif | Netral | Positif | Total |
|----------------------|---------|--------|---------|-------|
| Usability            | 164     | 112    | 3.467   | 3.743 |
| Fitur Aplikasi       | 292     | 133    | 1.895   | 2.320 |
| Keamanan             | 409     | 205    | 1.639   | 2.253 |
| Transaksi/Pembayaran | 187     | 173    | 992     | 1.352 |
| Customer Service     | 60      | 28     | 265     | 353   |

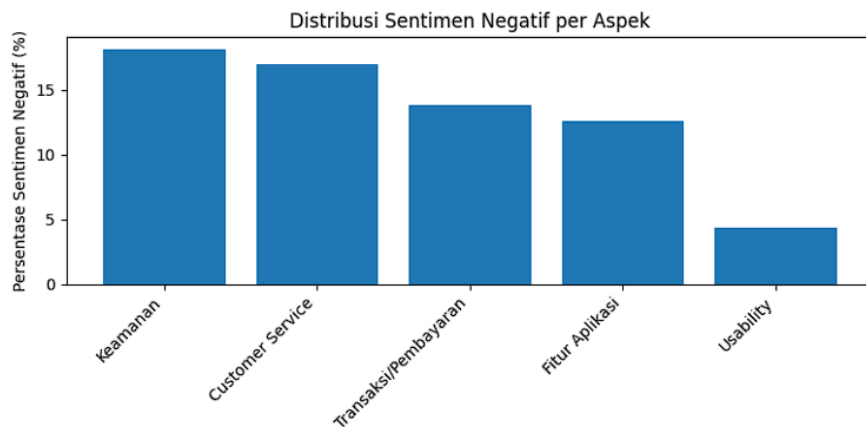
Distribusi sentimen positif, negatif, dan netral pada setiap aspek bisnis ditampilkan pada Tabel 8. Berdasarkan tabel distribusi sentimen per aspek, pada aspek Kemudahan Penggunaan terdapat 3.467 ulasan positif (92,63%), 164 ulasan negatif (4,38%), dan 112 ulasan netral (2,99%) dari total 3.743 ulasan. Aspek Fitur Aplikasi memiliki 1.895 ulasan positif (81,68%), 292 ulasan negatif (12,59%), dan 133 ulasan netral (5,73%) dari total 2.320 ulasan. Pada aspek Keamanan, terdapat 1.639 ulasan positif (72,77%), 409 ulasan negatif (18,15%), dan 205 ulasan netral (9,10 persen) dari total 2.253 ulasan.

Pada aspek Transaksi dan Pembayaran, terdapat 992 ulasan positif (73,37 %), 187 ulasan negatif (13,83 %), dan 173 ulasan netral (12,80 %) dari total 1.352 ulasan. Sementara itu, aspek Layanan Pelanggan memiliki 265 ulasan positif (75,07 %), 60 ulasan negatif (17,00 %), dan 28 ulasan netral (7,93 %) dari total 353 ulasan. Hasil ini menunjukkan bahwa meskipun sentimen positif ditemukan pada seluruh aspek, masih terdapat sejumlah ulasan negatif pada aspek Keamanan, Transaksi dan Pembayaran, serta Layanan Pelanggan.



Gambar 6. Heatmap Sentimen per Aspek

Visualisasi distribusi sentimen pada masing-masing aspek bisnis ditampilkan pada Gambar 6 dalam bentuk *heatmap*. *Heatmap* ini menunjukkan perbedaan distribusi sentimen positif, negatif, dan netral pada setiap aspek layanan aplikasi FLOQ, sehingga memudahkan identifikasi aspek yang memiliki kecenderungan sentimen negatif lebih tinggi dibandingkan aspek lainnya.



Gambar 7. Distribusi Sentimen Negatif Per Aspek

Distribusi sentimen negatif pada setiap aspek bisnis ditampilkan pada Gambar 7 untuk menyoroti aspek layanan yang paling sering dikaitkan dengan keluhan pengguna. Selain itu, distribusi sentimen negatif pada setiap aspek divisualisasikan menggunakan diagram batang. Visualisasi ini memperjelas perbedaan jumlah sentimen negatif antar aspek, di mana aspek Keamanan dan Layanan Pelanggan memiliki jumlah sentimen negatif yang relatif lebih tinggi dibandingkan aspek lainnya. Diagram ini membantu menyoroti aspek-aspek layanan yang paling sering menjadi sumber keluhan pengguna.

Secara keseluruhan, hasil penelitian ini menjawab rumusan masalah yang diajukan. Pertama, proses klasifikasi sentimen berhasil mengelompokkan ulasan pengguna aplikasi FLOQ ke dalam kategori positif, negatif, dan netral berdasarkan data ulasan yang dianalisis. Kedua, analisis distribusi sentimen menunjukkan variasi jumlah ulasan pada setiap kategori sentimen selama periode pengamatan, yang memberikan gambaran dinamika respons pengguna terhadap aplikasi. Ketiga, analisis aspek bisnis yang diterapkan pada lima aspek layanan yang telah ditetapkan dalam penelitian—Keamanan, Kemudahan Penggunaan, Fitur Aplikasi, Layanan Pelanggan, serta Transaksi/Pembayaran—menunjukkan perbedaan fokus pembahasan dan kecenderungan sentimen pada masing-masing aspek. Dengan demikian, kombinasi klasifikasi sentimen dan pemetaan aspek berbasis keyword matching mampu memberikan gambaran terstruktur mengenai persepsi pengguna terhadap berbagai dimensi layanan aplikasi FLOQ

## KESIMPULAN DAN SARAN

Berdasarkan eksperimen yang telah dilakukan terhadap 6.669 ulasan pengguna aplikasi FLOQ, diperoleh bahwa pembagian data latih dan data uji dengan rasio 70:30 menghasilkan kinerja model IndoBERT yang paling stabil dibandingkan rasio 80:20 dan 90:10. Dengan rasio ini, model mampu melakukan inferensi sentimen terhadap seluruh data ulasan dan menghasilkan 5.076 ulasan positif, 896 ulasan negatif, serta 697 ulasan netral. Hasil tersebut kemudian dianalisis secara temporal untuk melihat perubahan jumlah ulasan pada setiap periode pengamatan, serta dianalisis lebih lanjut menggunakan pendekatan aspek bisnis. Analisis aspek yang difokuskan pada lima aspek layanan, yaitu Keamanan, Kemudahan Penggunaan, Fitur Aplikasi, Layanan Pelanggan, serta Transaksi dan Pembayaran, menunjukkan bahwa aspek Kemudahan Penggunaan memiliki jumlah ulasan terbanyak, sementara sentimen negatif lebih sering muncul pada aspek Keamanan dan Layanan Pelanggan. Dengan demikian, tujuan penelitian untuk menganalisis sentimen pengguna, tren ulasan, serta persepsi pengguna berdasarkan aspek layanan aplikasi FLOQ telah tercapai sesuai dengan rumusan masalah yang ditetapkan.

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Data ulasan yang dianalisis hanya berasal dari satu platform aplikasi, sehingga hasil analisis belum sepenuhnya merepresentasikan persepsi seluruh pengguna FLOQ secara menyeluruh. Selain itu, pelabelan sentimen pada seluruh data dilakukan secara otomatis menggunakan model RoBERTa, sementara validasi manual hanya diterapkan pada sebagian data, sehingga masih dimungkinkan adanya ketidaktepatan label pada sebagian kecil ulasan. Di sisi lain, analisis aspek bisnis dilakukan menggunakan pendekatan berbasis aturan (*rule-based*) yang bergantung pada daftar kata kunci tertentu, sehingga belum sepenuhnya mampu menangkap aspek layanan yang disampaikan secara implisit oleh pengguna. Oleh karena itu, hasil penelitian ini perlu dipahami dengan mempertimbangkan keterbatasan data dan metode yang digunakan.

Penelitian selanjutnya dapat membandingkan performa model IndoBERT dengan model Transformer lain yang relevan untuk melihat perbedaan kinerja dalam klasifikasi sentimen, khususnya pada kelas sentimen netral yang masih memiliki potensi ambiguitas. Selain itu, penerapan pendekatan *Aspect-Based Sentiment Analysis* dengan metode yang lebih kontekstual, seperti berbasis pembelajaran mesin, dapat digunakan untuk memperoleh analisis sentimen yang lebih rinci pada setiap aspek layanan. Penelitian lanjutan juga dapat mempertimbangkan penggunaan data dengan rentang waktu yang lebih panjang atau sumber ulasan tambahan agar analisis tren dan persepsi pengguna dapat digambarkan secara lebih komprehensif.



## DAFTAR PUSTAKA

- [1] C. A. Ramaputra, M. H. Z. Al Faroby, and B. R. Lidiawaty, "Sentiment Analysis of User Reviews on Cryptocurrency Application: Evaluating the Impact of Dataset Split Scenarios Using Multinomial Naive Bayes," vol. 13, no. 4, pp. 4523–4535, 2024.
- [2] S. A. Helmayanti, F. Hamami, and R. Y. Fa'rifah, "PENERAPAN ALGORITMA TF-IDF DAN NAÏVE BAYES UNTUK ANALISIS SENTIMEN BERBASIS ASPEK ULASAN APLIKASI FLIP PADA GOOGLE PLAY STORE Sheva," *J. Indones. Manaj. Inform. dan Komun.*, vol. 4, no. 3, pp. 1822–1834, 2023.
- [3] K. L. Tan, C. P. Lee, and K. M. Lim, "A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research," *Appl. Sci.*, vol. 13, no. 7, 2023, doi: 10.3390/app13074550.
- [4] F. Y. A'la, "Indonesian Sentiment Analysis towards MyPertamina Application Reviews by Utilizing Machine Learning Algorithms," *J. Informatics Inf. Syst. Softw. Eng. Appl.*, vol. 5, no. 1, pp. 80–91, 2022, doi: 10.20895/inista.v5i1.838.
- [5] A. I. Hamid and A. M. Abdulazeez, "Sentiment Analysis Based on Machine Learning Techniques: A Comprehensive Review," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 4079–4106, 2024, [Online]. Available: <http://ijcs.stmikindonesia.ac.id/ijcs/index.php/ijcs/article/view/3135>
- [6] Siti Rahmawati, Dewi Anggraini, and Rizki Kurniawan, "Natural Language Processing For Automatic Sentiment Analysis In Social Media Data," *Int. J. Inf. Eng. Sci.*, vol. 1, no. 1, pp. 16–19, 2024, doi: 10.62951/ijies.v1i1.54.
- [7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, Association for Computational Linguistics, 2019.
- [8] A. Vaswani *et al.*, "Attention Is All You Need," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017.
- [9] B. Wilie *et al.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," in *Proceedings of the 28th International Conference on Computational Linguistics*, International Committee on Computational Linguistics, 2020, pp. 843–857.
- [10] A. Simanungkalit, J. P. P. Naibaho, and A. De Kweldju, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Shopee Menggunakan Algoritma Naïve Bayes," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, pp. 659–670, 2024.
- [11] A. Novrisal, G. I. Marthasari, and C. S. K. Aditya, "Sentimen Analisis Tweet Berbahasa Indonesia Pada Pilkada Serentak 2020 Menggunakan Metode Naïve Bayes Berbasis Particle Swarm Optimization Adam," vol. 3, no. 2, pp. 191–198, 2021.
- [12] N. A. Salsabila, "Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Gus Dur Menggunakan Metode Naïve Bayes Dan Support Vector Machine (SVM)," Universitas Islam Negeri Syarif Hidayatullah, 2022.
- [13] Dhendra Marutho and V. Gayuh Utomo, "Benchmarking IndoBERT and Transformer Models for Sentiment Classification on Indonesian E-Government Service Reviews," *J. Transform.*, vol. 23, no. 1, pp. 86–95, 2025, doi: 10.26623/transformatika.v23i1.12095.
- [14] Khen Dedes, Fatimatu Zahra, M. Hermansyah, A. B. Setiawan, R. P. Pradana, and A. F. M. Harvyanti, "BERT Sentimen: Fine-Tuning Multibahasa untuk Ulasan Bahasa Indonesia," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 2, pp. 1080–1084, 2025, doi: 10.62712/juktisi.v4i2.585.
- [15] R. L. Mustofa, Tarno, and C. E. Widodo, "Analisis Sentimen Berbasis Aspek Pada Aplikasi Elektronik Survei Kepuasan Masyarakat (E-Skm) Jawa Tengah Menggunakan Indobert Aspect-Based Sentiment Analysis on the Central Java Public Satisfaction Survey Electronic Application (E-Skm) Using Indobert," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 4, pp. 867–874, 2025.
- [16] P. Khabiibah, S. W. Sihwi, and H. Setiadi, "Aspect Based Opinion Summarization Using Rule-Based Method and Support Vector Machine for Indonesian Reviews," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 12, no. 4, pp. 1379–1386, 2022, doi: 10.18517/ijaseit.12.4.12716.
- [17] S. P. Mahardika, R. Mardhiyyah, and F. I. Sanjaya, "ANALISIS SENTIMEN PUBLIK TERHADAP BADAN INVESTASI DANANTARA PADA MEDIA SOSIAL X MENGGUNAKAN MODEL INDOBERT," vol. 7, no. 4, pp. 1711–1721, 2025.
- [18] S. M. A. Zahra, "Analisis Sentimen Pengguna Terhadap Aplikasi Telemedicine Menggunakan Metode Fine-Tuning IndoBERT," Universitas Islam Negeri Maulana Malik Ibrahim Malang, 2025.
- [19] L. S. Pradana, "Analisis Sentimen Masyarakat Media Sosial Twitter terhadap Kinerja Penjabat Gubernur DKI Jakarta menggunakan Model INDOBERT," Universitas Islam Negeri Syarif Hidayatullah Jakarta, 2024.
- [20] M. P. Firdaus and D. Trisnawarman, "Analisis Sentimen Publik Terhadap Program Tabungan Perumahan Rakyat Menggunakan Model IndoBERT Lite pada Komentar YouTube," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 1, pp. 359–368, 2025.
- [21] U. Khairani, V. Mutiawani, and H. Ahmadian, "PENGARUH TAHAPAN PREPROCESSING TERHADAP MODEL INDOBERT DAN THE INFLUENCE OF PREPROCESSING STAGES ON INDOBERT AND INDOBERTWEET MODELS FOR EMOTION DETECTION IN INSTAGRAM NEWS," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 4, pp. 887–894, 2024, doi: 10.25126/jtiik.1148315.
- [22] M. A. Palimbani, R. P. Hasuti, and R. A. Rajagade, "Analisis Sentimen Berbasis Aspek pada Ulasan Pengguna Aplikasi Starbucks Menggunakan Algoritma Support Vector Machine," vol. 5, no. 1, pp. 43–49, 2024.
- [23] A. AlMasaud and H. H. Al-Baity, "AraMAMS: Arabic Multi-Aspect, Multi-Sentiment Restaurants Reviews Corpus for Aspect-Based Sentiment Analysis," *Sustainability*, vol. 15, no. 16, 2023, doi: 10.3390/su151612268.
- [24] F. L. Hakim and K. E. Dewi, "ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP PRODUK KECANTIKAN MENGGUNAKAN NEIGHBOR WEIGHTED K-NEAREST NEIGHBOR," *KOMPUTA J. Ilm. Komput. dan Inform.*, vol. 13, no. 1, pp. 1–10, 2024.
- [25] M. Sivakumar, S. Parthasarathy, and T. Padmapriya, "Trade-off between training and testing ratio in machine learning for medical image processing," *PeerJ Comput. Sci.*, 2024, doi: 10.7717/peerj-cs.2245.