

Sentiment Analysis on BRImo Application Reviews Using IndoBERT

Asyer Aprinando Pratama Simarmata ^{1*}, Theopilus Bayu Sasongko ^{2*}

* Informatika, Universitas Amikom Yogyakarta

asyersimarmata@students.amikom.ac.id¹, theopilus.27@amikom.ac.id²

Article Info

Article history:

Received 2024-08-01

Revised 2025-03-17

Accepted 2025-03-26

Keyword:

BRImo,

Digital Banking,

IndoBERT,

Natural Language Processing,

Sentiment Analysis.

ABSTRACT

The advancement of information technology has significantly impacted various sectors, including digital banking. BRImo, a mobile banking application from Bank Rakyat Indonesia (BRI), has been widely used, generating numerous user reviews that reflect their experiences. This study applies IndoBERT, a transformer-based model specifically designed for the Indonesian language, to analyze sentiment in BRImo user reviews. IndoBERT excels in handling the unique characteristics of the Indonesian language, such as informal and mixed-language usage. The dataset was collected from Kaggle and processed through labeling, data balancing, and splitting into 80% training, 10% validation, and 10% testing data. The IndoBERT model was evaluated using a confusion matrix and achieved 90% accuracy, with F1-scores of 0.89 for negative, 0.91 for neutral, and 0.90 for positive sentiments. Sentiment analysis results indicate that a significant portion of negative reviews highlight issues related to login difficulties, transaction failures, and slow customer service response times. These insights can help BRI enhance application reliability and customer support efficiency. This study demonstrates that IndoBERT is effective in sentiment analysis for Indonesian text and can be utilized to enhance BRImo services by providing deeper insights into user feedback.



This is an open access article under the [CC-BY-SA](#) license.

I. PENDAHULUAN

Kemajuan teknologi informasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk sektor perbankan. Salah satu inovasi penting dalam bidang ini adalah munculnya aplikasi perbankan digital yang memudahkan nasabah dalam melakukan berbagai transaksi keuangan. Aplikasi *mobile banking* telah menjadi solusi praktis bagi nasabah untuk melakukan transaksi perbankan kapan saja dan di mana saja. Aplikasi perbankan digital yang dimiliki Bank Rakyat Indonesia (BRI) adalah BRImo. BRImo merupakan pembaharuan dari aplikasi *mobile banking* Bank BRI sebelumnya yaitu BRI Mobile yang masih mengandalkan penggunaan pesan singkat (*SMS Mobile*) untuk keperluan verifikasi [1]. Seiring dengan meningkatnya penggunaan aplikasi ini, ulasan dari pengguna juga semakin banyak. Ulasan-ulasan ini mencakup berbagai macam sentimen, mulai dari positif, netral, hingga negatif, yang semuanya penting bagi pengembang aplikasi untuk memahami pengalaman pengguna dan melakukan perbaikan yang diperlukan. Ulasan-ulasan tersebut dapat menjadi sumber data untuk dianalisa

dan diklasifikasi menggunakan analisis sentimen [2]. Sayangnya, dengan banyaknya ulasan yang masuk, proses manual untuk menganalisis sentimen ini menjadi tidak efisien dan memakan waktu. Oleh karena itu, diperlukan suatu metode otomatis untuk melakukan analisis sentimen terhadap ulasan-ulasan tersebut.

Analisis sentimen adalah proses otomatisasi pengolahan data teks untuk mengidentifikasi dan mengkategorikan opini yang diekspresikan dalam teks, khususnya untuk menentukan apakah sikap penulis terhadap topik tertentu positif, negatif, atau netral. Analisis sentimen dilakukan untuk mengidentifikasi sentimen atau emosi yang terkandung dalam teks, seperti ulasan pengguna di *Google Play Store* sehingga bisa memberikan informasi tentang aplikasi tersebut [3]. Penelitian ini berfokus pada bagaimana mengekstraksi dan memahami sentimen yang terkandung dalam ulasan-ulasan pengguna aplikasi BRImo. Analisis sentimen ini penting dilakukan untuk memperoleh wawasan yang mendalam mengenai persepsi pengguna terhadap aplikasi BRImo, yang nantinya dapat digunakan untuk meningkatkan kualitas layanan dan fitur aplikasi tersebut.

Penelitian terkait analisis sentimen yang menggunakan *IndoBERT* masih terbatas. Selain itu, aplikasi perbankan digital memiliki karakteristik ulasan pengguna yang unik, sehingga memerlukan pendekatan khusus dalam analisis sentimennya. Permasalahan yang akan dibahas dalam penelitian ini adalah bagaimana memanfaatkan *IndoBERT* untuk melakukan analisis sentimen pada ulasan aplikasi BRImo dengan akurasi yang tinggi. Banyak ulasan pengguna aplikasi perbankan digital yang mengandung bahasa informal, campuran bahasa, serta istilah-istilah teknis yang spesifik. Hal ini menambah kompleksitas dalam analisis sentimen, sehingga model yang digunakan harus mampu menangani berbagai variasi bahasa dan konteks yang berbeda. Selain itu, model harus dapat mengidentifikasi sentimen yang terkandung dalam ulasan-ulasan tersebut dengan tepat, agar pihak pengembang aplikasi dapat mengambil langkah-langkah yang sesuai untuk meningkatkan kualitas layanan.

Penelitian terdahulu yang dilakukan oleh Susandri dkk membandingkan kinerja *Support Vector Machine*, *Naïve Bayes*, dan *Adaboost* pada sentimen pengguna aplikasi BRImo, mendapatkan hasil yang menunjukkan *SVM* memiliki akurasi tertinggi sebesar 90,4%, *precision* 90,8%, *recall* 90%, dan *F1-score* 90,3% [4]. Penelitian yang telah dilakukan oleh Arminda dkk tentang implementasi *Multinomial Naïve Bayes* pada analisis sentimen ulasan pengguna aplikasi BRImo menggunakan *10-fold cross validation* mendapatkan hasil terbaik pada fold-2 dengan nilai akurasi 98,02%, presisi 97,06%, *recall* 97,06%, dan *f1-score* 97,06% [5]. Kemudian penelitian yang telah dilakukan oleh Khoirul dkk pada tahun 2023 tentang analisis sentimen aplikasi BRImo pada ulasan pengguna di *Google Play Store* dengan menggunakan *Naïve Bayes* mendapatkan tingkat *accuracy* sebesar 84,52% *precision* 82,51% dan *recall* 87,62% [6]. Umair dan Sutanto juga telah melakukan penelitian analisis sentimen ulasan pengguna aplikasi BRImo menggunakan *Naïve Bayes* pada tahun 2024 dengan hasil nilai akurasi 65%, nilai presisi 67%, nilai *recall* 92% dan *f1_score* 77% [7]. Penelitian lain yang dilakukan oleh Nuraini dkk mengenai analisis sentimen terhadap *review* aplikasi BRImo di *Google Play Store* menggunakan *Naïve Bayes* mendapatkan hasil akurasi sebesar 87.92%, presisi sebesar 83.65% *recall* sebesar 87.92% dan *F1 Score* sebesar 84.39% [8].

Dalam penelitian ini, peneliti menggunakan *IndoBERT*, sebuah model bahasa berbasis *transformer* yang telah dilatih secara ekstensif pada data teks bahasa Indonesia, untuk melakukan analisis sentimen pada ulasan aplikasi BRImo. *IndoBERT* memiliki desain arsitektur yang mirip dengan BERT, tetapi dikondisikan pada Indo4B dalam dua fase dengan kumpulan 4 miliar kata dan sekitar 250 juta kalimat [9]. *IndoBERT* memiliki keunggulan dalam memahami konteks bahasa Indonesia dibandingkan model-model sebelumnya, sehingga diharapkan dapat memberikan hasil yang lebih akurat dan relevan. Model ini telah dilatih pada korpus besar teks bahasa Indonesia, sehingga mampu mengenali konteks bahasa secara lebih akurat, terutama dalam menghadapi bahasa informal dan campuran. Selain itu,

penelitian sebelumnya masih terbatas dalam membahas implementasi *IndoBERT* untuk analisis sentimen dalam sektor perbankan digital. Penelitian terkait yang membandingkan metode seperti *Naïve Bayes* dan *SVM* telah menunjukkan hasil yang bervariasi, dengan akurasi berkisar antara 65% hingga 98%, tergantung pada teknik *pre-processing* dan fitur yang digunakan. Seperti pada penelitian Tarwoto dkk yang menunjukkan keunggulan *IndoBERT* dalam memahami makna dan konteks serta pola yang kompleks pada teks bahasa Indonesia dibandingkan dengan *SVM* dan *Naïve Bayes* [10]. Penelitian Hakim dkk juga mengindikasikan bahwa *IndoBERT* yang merupakan model monolingual lebih unggul dibandingkan model multilingual seperti *MultilingualBERT* dalam menangani tugas bahasa spesifik, karena akan lebih fokus dalam memahami struktur dan nuansa satu bahasa secara mendalam [11].

Metode penelitian yang digunakan meliputi pengumpulan data ulasan aplikasi BRImo dari platform *Kaggle*, pra-pemrosesan data untuk membersihkan dan mempersiapkan teks ulasan, serta penerapan model *IndoBERT* untuk analisis sentimen. Data ulasan yang dikumpulkan akan melalui tahap pra-pemrosesan seperti penghapusan karakter spesial dan tokenisasi. Selanjutnya, model *IndoBERT* yang telah dilatih akan digunakan untuk mengklasifikasikan sentimen dalam ulasan tersebut ke dalam kategori positif, netral, atau negatif.

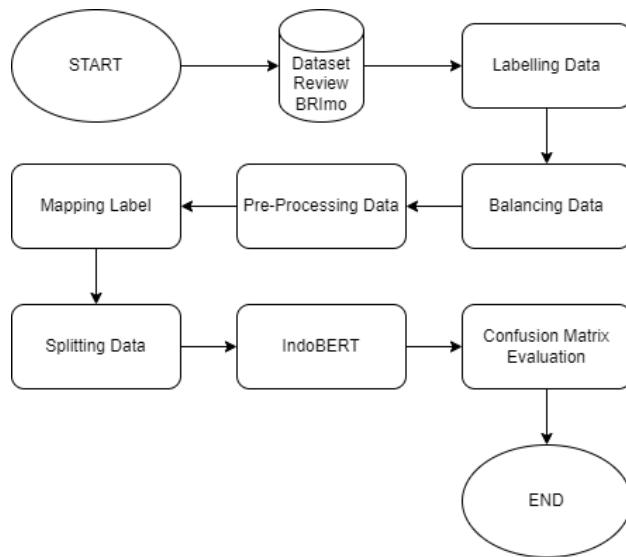
Penelitian ini bertujuan untuk mengembangkan model analisis sentimen menggunakan algoritma *IndoBERT* untuk mengklasifikasikan ulasan-ulasan pada aplikasi BRImo. Selain itu, penelitian ini juga bertujuan untuk mengidentifikasi dan mengatasi tantangan dalam analisis sentimen ulasan aplikasi perbankan digital yang disesuaikan dengan karakteristik ulasan aplikasi BRImo dan memberikan rekomendasi untuk peningkatan kualitas layanan berdasarkan hasil analisis sentimen ulasan pengguna. Dengan mencapai tujuan-tujuan ini, diharapkan penelitian ini dapat memberikan kontribusi signifikan dalam bidang analisis sentimen berbasis bahasa Indonesia serta peningkatan layanan aplikasi perbankan digital.

Manfaat yang diharapkan dari hasil penelitian ini tidak hanya bermanfaat bagi BRI dalam meningkatkan layanan aplikasi BRImo, tetapi juga memberikan kontribusi dalam pengembangan metode analisis sentimen yang lebih efektif khususnya untuk bahasa Indonesia, serta implementasi *IndoBERT* dalam tugas-tugas NLP di Indonesia. Dengan memanfaatkan kekuatan *IndoBERT* dalam memahami teks, penelitian ini juga diharapkan dapat membuka peluang untuk penerapan lebih luas dalam berbagai domain yang memerlukan analisis sentimen, seperti pemasaran, layanan pelanggan, dan evaluasi produk.

II. METODE

Alur penelitian memuat *flowchart* langkah-langkah penelitian secara lengkap dan rinci [12]. Penelitian ini dimulai dengan mengunduh *dataset* ulasan aplikasi BRImo dari *Kaggle*, dilanjutkan dengan pelabelan, penyeimbangan, dan pra-pemrosesan data. Data yang telah diproses kemudian

dipetakan dan dibagi menjadi set pelatihan, validasi dan pengujian. Kemudian model *IndoBERT* dilatih, dan kinerjanya dievaluasi menggunakan *confusion matrix*. Alur penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

A. Unduh Dataset

Tahap pertama dalam penelitian ini adalah pengumpulan dataset, dengan cara melakukan pengunduhan dataset ulasan aplikasi BRImo dari *Kaggle*. *Kaggle* adalah platform *data science* online yang menyediakan kompetisi *data science* untuk masalah bisnis, perekrutan, dan tujuan penelitian akademis [13]. *Kaggle* merupakan platform penyedia dataset yang sering digunakan oleh peneliti dan praktisi data. Dataset ini merupakan kumpulan ulasan dari pengguna aplikasi BRImo yang diambil dari platform *Google Play Store*. Setiap ulasan akan mencakup teks ulasan serta informasi tambahan seperti *rating score* yang diberikan oleh pengguna.

B. Labelling Data

Setelah dataset diunduh dan dimuat di *Google Colab*, langkah berikutnya adalah pemberian label pada data. *Labelling* data diperlukan untuk menentukan kelas dari setiap ulasan [14]. Setiap ulasan akan dikategorikan ke dalam tiga kelas sentimen: positif, negatif, atau netral, yang ditentukan berdasarkan *rating score*. Sentimen positif untuk *score* 4 dan 5, sentimen netral untuk *score* 3, dan sentimen negatif untuk *score* 1 dan 2. Label ini sangat penting untuk melatih model analisis sentimen yang akan digunakan.

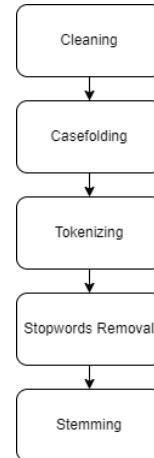
C. Balancing Data

Tahap ini bertujuan untuk memastikan distribusi kelas sentimen dalam dataset seimbang. Dalam banyak kasus, dataset ulasan seringkali tidak seimbang, dimana kelas tertentu mungkin lebih dominan dibandingkan kelas lainnya. Dalam penelitian ini sendiri dataset yang digunakan mengalami ketidakseimbangan kelas, dengan lebih banyak

ulasan positif dibandingkan netral dan negatif. Untuk mengatasi hal ini, teknik penyeimbangan data yang digunakan adalah *random undersampling* pada kelas mayoritas (positif) dan *random oversampling* pada kelas minoritas (netral) agar distribusi lebih seimbang. *Random undersampling* akan mengurangi kelas mayoritas dengan menghilangkan sampel secara acak sehingga jumlah sampel di kelas mayoritas sama dengan kelas minoritas, sedangkan *random oversampling* akan meningkatkan kelas minoritas dengan menduplikasi sampel secara acak sehingga jumlah sampel di kelas minoritas sama dengan kelas mayoritas [15]. Penyeimbangan ini penting untuk mencegah model menjadi bias terhadap salah satu kategori/kelas.

D. Pre-Processing Data

Pre-processing data adalah tahapan penting dalam analisis data, yang melibatkan persiapan data mentah dan mengubahnya menjadi format yang lebih terstruktur dan siap untuk dianalisis [16]. Proses ini bertujuan tidak hanya untuk meningkatkan kualitas data, namun juga membuat data siap digunakan secara efektif oleh model [16]. Pra-pemrosesan data melibatkan beberapa tahapan untuk membersihkan dan mempersiapkan teks ulasan. Tahap-tahap ini dapat dilihat pada Gambar 2.



Gambar 2. Tahapan Pre-Processing Data

1) *Cleaning*: Merupakan tahap awal dalam pra-pemrosesan data yang bertujuan untuk membersihkan teks dari karakter-karakter yang tidak relevan atau mengganggu proses analisis. Pada tahap ini, tanda baca seperti titik, koma, tanda seru, dan tanda tanya dihapus [17]. Selain itu, karakter spesial seperti simbol @, #, \$, %, dan sebagainya juga dihilangkan. Angka yang tidak relevan dengan analisis sentimen dihapus, begitu juga dengan spasi berlebih dan karakter non-alfabet lainnya.

2) *Casefolding*: Dalam langkah ini, seluruh teks dikonversi menjadi huruf kecil [18] guna mencapai konsistensi dan mengurangi variasi yang tidak diperlukan. Hal ini membantu menghindari kesalahan dalam analisis yang

disebabkan oleh perbedaan huruf besar dan kecil, sehingga teks menjadi lebih homogen dan mudah diolah oleh model.

3) *Tokenizing*: Proses yang memisahkan kata dari kata lain [19]. Proses ini memecah teks menjadi unit-unit kata yang lebih kecil yang disebut token, yang sangat penting untuk analisis lebih lanjut. Langkah ini melibatkan pemisahan teks berdasarkan spasi dan tanda baca untuk mendapatkan kata-kata individual. Dengan melakukan tokenisasi, teks menjadi lebih terstruktur dan memudahkan proses analisis karena model *NLP* bekerja lebih baik dengan kata-kata yang terpisah.

4) *Stopwords Removal*: Bertujuan untuk menghilangkan kata-kata umum dalam kalimat atau kata tertentu [20] yang tidak memiliki makna signifikan dalam konteks analisis sentimen. Pada tahap ini, kata-kata seperti "dan", "atau", "dengan", "yang", dan lain-lain yang sering muncul tetapi tidak membawa makna penting diidentifikasi dan dihapus. Dengan menghapus *stopwords*, analisis menjadi lebih efisien karena fokus diberikan pada kata-kata yang membawa informasi lebih penting.

5) *Stemming*: Proses mengubah kata-kata ke bentuk dasarnya atau menghilangkan imbuhan [21] untuk menyatukan variasi kata yang memiliki arti serupa. Langkah ini menggunakan *library stemming Sastrawi* untuk mengubah kata-kata menjadi bentuk dasar atau akar katanya. Proses *stemming* mengurangi kompleksitas dan variasi dalam data teks, sehingga model dapat lebih mudah mengenali dan mempelajari pola dari data tersebut, meningkatkan akurasi dan efektivitas analisis sentimen.

E. Mapping Label

Setelah *pre-processing* data dilakukan, langkah berikutnya adalah memetakan label. Proses ini memetakan kolom label sentimen ke dalam bentuk numerik. Label sentimen negatif = 0, netral = 1, dan positif = 2.

F. Splitting Data

Pada proses ini, *dataset* yang telah dipetakan kemudian akan dibagi menjadi tiga set: data pelatihan (*training set*), data validasi (*validation set*) dan data pengujian (*testing set*) [22]. Pembagian dilakukan dengan rasio 80:10:10, di mana sebagian besar data digunakan untuk melatih model, sementara sisanya digunakan untuk menguji kinerja model. Pada penelitian ini dataset dibagi dengan rasio 80:10:10 dikarenakan hasil yang jauh lebih baik daripada rasio 70:15:15 seperti yang ditunjukkan di penelitian yang dilakukan oleh Sumalatha dkk [23].

G. IndoBERT

IndoBERT merupakan varian Indonesia dari model *pre-trained BERT*, digunakan untuk melakukan analisis sentimen pada dataset teks bahasa Indonesia [24]. Pada tahap ini, model *IndoBERT* dilatih menggunakan data pelatihan yang telah diproses. Proses pelatihan melibatkan pengaturan *hyperparameter max length, batch size, learning rate* dan *epoch* untuk meningkatkan performa model [25]:

1) *Max Length*: Panjang maksimal token yang digunakan untuk setiap ulasan.

2) *Batch Size*: Jumlah sampel yang digunakan untuk *backward* dan *forward pass* [26], yang diproses dalam satu iterasi pelatihan.

3) *Learning Rate*: Kecepatan pembaruan bobot model selama pelatihan.

4) *Epochs*: Jumlah iterasi pelatihan dilakukan menggunakan data pelatihan [26].

Model dilatih untuk mengenali pola dan karakteristik dalam data ulasan yang dapat digunakan untuk menentukan sentimen setiap ulasan.

H. Confusion Matrix Evaluation

Setelah model dilatih, langkah berikutnya adalah evaluasi kinerja model menggunakan data pengujian. Kinerja model diuji menggunakan data pengujian untuk mengevaluasi kemampuan model dalam mengklasifikasikan sentimen ulasan. Hasil pengujian dianalisis menggunakan *confusion matrix* untuk mengidentifikasi kekuatan dan kelemahan model. Evaluasi dilakukan dengan menggunakan metrik-metrik evaluasi seperti *accuracy, precision, recall, dan f1-score* [27], yang membantu dalam menilai performa model.

1) *Accuracy*: Perbandingan antara jumlah prediksi yang benar dengan total jumlah prediksi. *Accuracy* adalah ukuran yang menunjukkan seberapa sering model membuat prediksi yang benar. Rumusnya adalah:

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

2) *Precision*: Ukuran yang menunjukkan seberapa banyak prediksi positif yang benar dibandingkan dengan seluruh prediksi positif yang dibuat oleh model. *Precision* memberikan gambaran tentang akurasi dari prediksi positif. Rumusnya adalah:

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

3) *Recall*: Ukuran yang menunjukkan seberapa banyak prediksi positif yang benar dibandingkan dengan seluruh data yang sebenarnya positif. *Recall* memberikan gambaran tentang kemampuan model dalam menangkap semua data positif. Rumusnya adalah:

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

4) *F1-Score*: Ukuran harmonis dari precision dan recall. *F1-Score* adalah metrik yang memperhitungkan *precision* dan *recall*. *F1-Score* memberikan gambaran tentang kinerja keseluruhan dari model. Rumusnya adalah:

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

III. HASIL DAN PEMBAHASAN

Peneliti menggunakan *dataset* yang diunduh dari *Kaggle*. *Dataset* ini berisi ulasan aplikasi BRImo dari *Google Play Store*, yang memiliki 17219 baris dan 9 kolom [28]. *Dataset* kemudian dimuat ke *Google Colab* sebagai *DataFrame* awal, yang dapat dilihat pada Tabel I.

TABEL I
DATAFRAME AWAL (5 BARIS TERATAS)

Unnamed: 0	review Id	content	score	thumbsUp Count	reviewCreated Version	at	replyContent	replied At
0	c4f795 94- 78cd- 48d5- 876f- ef864d 13657a	Memudahkan saya untuk melakukan transaksi apapun	5	0	2.35.0	1/31/20 23 23:59		
1	88754d f5- 86f1- 42f7- ba5d- a9ca5d f6c1e6	Mantap	5	0	2.35.0	1/31/20 23 23:59		
2	8f5ca7 92- d4f2- 4ea9- 9c9f- 6720bc 3e942c	Aplikasi nya sangat membantu segala bentuk transaksi saya mantap	5	0		1/31/20 23 23:59		
3	0297f1 bd- 2dc9- 4509- 9c72- f18019 64fb87	simpel cepat. ga repot	5	0	2.31.0	1/31/20 23 23:58		
4	72617e 6a- 048b- 42cd- bce3- c526a4 afe9a5	bagus	5	0	2.35.0	1/31/20 23 23:55		

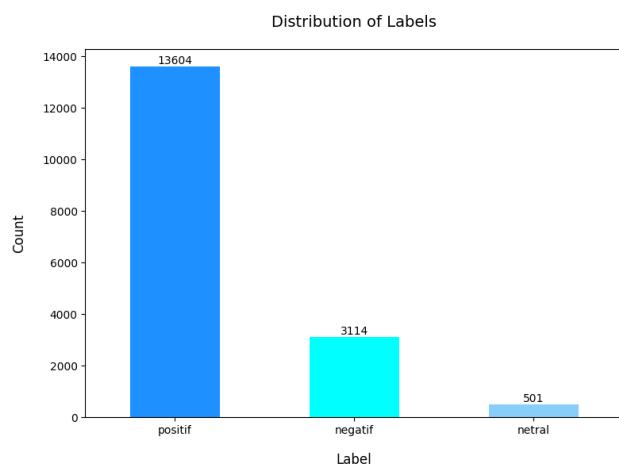
Dari *DataFrame* awal dipilih dua kolom yang relevan, yaitu kolom *content*, yang berisi teks ulasan dari pengguna aplikasi, dan kolom *score*, yang berisi *rating* nilai 1 sampai 5. Setelah itu peneliti menambahkan kolom baru, yaitu ‘label’. Label diberikan berdasarkan kolom *score*. Label ‘positif’ untuk *score* 4 dan 5, label ‘netral’ untuk *score* 3, dan label

‘negatif’ untuk *score* 1 dan 2. Ketiga kolom tersebut lalu disimpan menjadi sebuah *DataFrame* baru, yang dapat dilihat pada Tabel II.

TABEL II
DATAFRAME BARU (5 BARIS TERATAS)

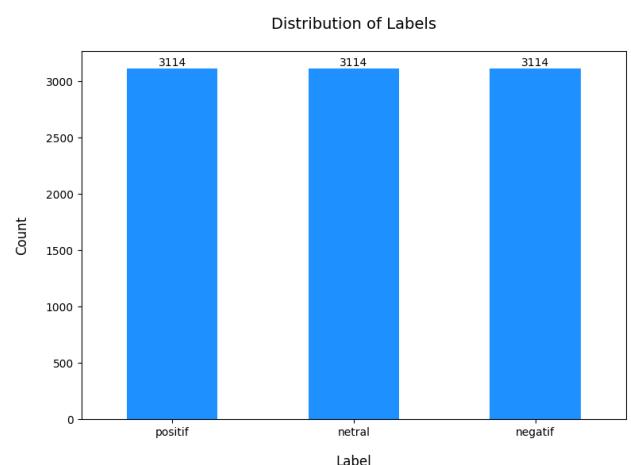
content	score	label
Memudahkan saya untuk melakukan transaksi apapun	5	positif
Mantap	5	positif
Aplikasi nya sangat membantu segala bentuk transaksi saya mantap	5	positif
simpel cepat. ga repot	5	positif
bagus	5	positif

Setelah proses *labelling*, peneliti mengecek apakah ada nilai *null* dan apakah *DataFrame* seimbang dalam setiap kategorinya. Meskipun tidak ditemukan nilai *null*, namun didapatkan ketidakseimbangan dalam pendistribusian label dengan rincian 13604 untuk label positif, 3114 untuk label negatif, dan 501 untuk label netral. Distribusi label dapat dilihat pada Gambar 3.



Gambar 3. Distribusi Label

Berdasarkan ketidakseimbangan ini kemudian dilakukan *balancing data* (penyeimbangan data). Peneliti memutuskan untuk melakukan *balancing* dengan cara *undersampling* label mayoritas (positif) dan *oversampling* label minoritas (netral) untuk menyamai label negatif yang berjumlah 3114. *Undersampling* label positif dilakukan dengan mengambil hanya 3114 sampel secara acak tanpa penggantian. *Oversampling* label netral dilakukan dengan mengambil 3114 sampel dengan penggantian (*sampling with replacement*). *Oversampling* yang dilakukan menghasilkan sampel baru dari data yang sudah ada dalam kategori netral. Dalam konteks ini, beberapa sampel yang sudah ada dipilih secara acak untuk dimasukkan kembali ke *dataset* dengan tujuan meningkatkan jumlah sampel hingga mencapai jumlah yang diinginkan. Setelah melakukan *undersampling* dan *oversampling*, ketiga *dataset* yang telah seimbang digabungkan menjadi satu *dataset*. Kemudian *dataset* ini diacak untuk memastikan urutan data acak sebelum dilakukan analisis atau pemodelan lebih lanjut. Hasil penyeimbangan data dapat dilihat pada Gambar 4.

Gambar 4. Distribusi Label Setelah *Balancing*

DataFrame setelah *balancing* dapat dilihat pada Tabel III.

TABEL III
DATAFRAME BARU SETELAH BALANCING (5 BARIS TERATAS)

content	score	label
Belum pernah	5	positif
Saya sdh upgrade tp pas di buka kbagus gak bisa	3	netral
Sangat ribet ,pertama saya kebank untuk pergantian kartu atm karna rusak ,kata pihak bank kartu terbatas hanya untuk yg sudah tua yg usia muda direkomendasikan menggunakan brimo, trus langsung daftar dics karna dari hp gagal terus dan ga langsung jadi katanya harus datang lagi lain waktu saya bilang saya ada waktu pas libur kerja hari ini aja dan nunggu lagi libur 2mnggu kemudian , dan tetap ga bisa hari itu juga , skrng diulang ulang daftar via hp gagal terus ,kalo bisa ribet knapa harus mudah	1	negatif
Biasanya bisa dipakai untuk transaksi belanja online Lazada khusus mode transfer BNI dan Bank lainnya, tapi sekarang tidak bisa lagi tambah susah... Tolong solusinya dong ..	3	netral
bagus	5	positif

Tahap selanjutnya adalah *pre-processing data*. Langkah-langkah yang dilakukan adalah *cleaning*, *casefolding*, *tokenizing*, *stopwords removal*, dan *stemming*. Hasil *cleaning* dapat dilihat pada Tabel IV.

TABEL IV
HASIL CLEANING

Sebelum Cleaning	Setelah Cleaning
Belum pernah	Belum pernah
Saya sdh upgrade tp pas di buka kbagus gak bisa	Saya sdh upgrade tp pas di buka kbagus gak bisa
Sangat ribet ,pertama saya kebank untuk pergantian kartu atm karna rusak ,kata pihak bank kartu terbatas hanya untuk yg sudah tua yg usia muda direkomendasikan menggunakan brimo, trus langsung daftar dics karna dari hp gagal terus dan ga langsung jadi katanya harus datang lagi lain waktu saya bilang saya ada waktu pas libur kerja hari ini aja dan nunggu lagi libur 2mnggu kemudian , dan tetap ga bisa hari itu juga , skrng diulang ulang daftar via hp gagal terus ,kalo bisa ribet knapa harus mudah	Sangat ribet pertama saya kebank untuk pergantian kartu atm karna rusak kata pihak bank kartu terbatas hanya untuk yg sudah tua yg usia muda direkomendasikan menggunakan brimo trus langsung daftar dics karna dari hp gagal terus dan ga langsung jadi katanya harus datang lagi lain waktu saya bilang saya ada waktu pas libur kerja hari ini aja dan nunggu lagi libur mnggu kemudian dan tetap ga bisa hari itu juga skrng diulang ulang daftar via hp gagal terus kalo bisa ribet knapa harus mudah
Biasanya bisa dipakai untuk transaksi belanja online Lazada khusus mode transfer BNI dan Bank lainnya, tapi sekarang tidak bisa lagi tambah susah... Tolong solusinya dong .. bagus	Biasanya bisa dipakai untuk transaksi belanja online Lazada khusus mode transfer BNI dan Bank lainnya tapi sekarang tidak bisa lagi tambah susah Tolong solusinya dong .. bagus

Hasil *casefolding* dapat dilihat pada Tabel V.

TABEL V
HASIL CASEFOLDING

Sebelum Casefolding	Setelah Casefolding
Belum pernah	belum pernah
Saya sdh upgrade tp pas di buka kbagus gak bisa	saya sdh upgrade tp pas di buka kbagus gak bisa
Sangat ribet pertama saya kebank untuk pergantian kartu atm karna rusak kata pihak bank kartu terbatas hanya untuk yg sudah tua yg usia muda direkomendasikan menggunakan brimo trus langsung daftar dics karna dari hp gagal terus dan ga langsung jadi katanya harus datang lagi lain waktu saya bilang saya ada waktu pas libur kerja hari ini aja dan nunggu lagi libur mnggu kemudian dan tetap ga bisa hari itu juga skrng diulang ulang daftar via hp gagal terus kalo bisa ribet knapa harus mudah	sangat ribet pertama saya kebank untuk pergantian kartu atm karna rusak kata pihak bank kartu terbatas hanya untuk yg sudah tua yg usia muda direkomendasikan menggunakan brimo trus langsung daftar dics karna dari hp gagal terus dan ga langsung jadi katanya harus datang lagi lain waktu saya bilang saya ada waktu pas libur kerja hari ini aja dan nunggu lagi libur mnggu kemudian dan tetap ga bisa hari itu juga skrng diulang ulang daftar via hp gagal terus kalo bisa ribet knapa harus mudah

Biasanya bisa dipakai untuk transaksi belanja online Lazada khusus mode transfer BNI dan Bank lainnya tapi sekarang tidak bisa lagi tambah susah Tolong solusinya dong .. bagus	biasanya bisa dipakai untuk transaksi belanja online lazada khusus mode transfer bni dan bank lainnya tapi sekarang tidak bisa lagi tambah susah tolong solusinya dong .. bagus
---	---

Hasil *tokenizing* dapat dilihat pada Tabel VI.

TABEL VI
HASIL TOKENIZING

Sebelum Tokenizing	Setelah Tokenizing
belum pernah	[belum, pernah]
saya sdh upgrade tp pas di buka kbagus gak bisa	[saya, sdh, upgrade, tp, pas, di, buka, kbagus, gak, bisa]
sangat ribet pertama saya kebank untuk pergantian kartu atm karna rusak kata pihak bank kartu terbatas hanya untuk yg sudah tua yg usia muda direkomendasikan menggunakan brimo trus langsung daftar dics karna dari hp gagal terus dan ga langsung jadi katanya harus datang lagi lain waktu saya bilang saya ada waktu pas libur kerja hari ini aja dan nunggu lagi libur mnggu kemudian dan tetap ga bisa hari itu juga skrng diulang ulang daftar via hp gagal terus kalo bisa ribet knapa harus mudah	[sangat, ribet, pertama, saya, kebank, untuk, pergantian, kartu, atm, karna, rusak, kata, pihak, bank, kartu, terbatas, hanya, untuk, yg, sudah, tua, yg, usia, muda, direkomendasikan, menggunakan, brimo, trus, langsung, daftar, dics, karna, dari, hp, gagal, terus, dan, ga, langsung, jadi, katanya, harus, datang, lagi, lain, waktu, saya, bilang, saya, ada, waktu, pas, libur, kerja, hari, ini, aja, dan, nunggu, lagi, libur, mnggu, kemudian, dan, tetap, ga, bisa, hari, itu, juga, skrng, diulang, ulang, daftar, via, hp, gagal, terus, kalo, bisa, ribet, knapa, harus, mudah]
biasanya bisa dipakai untuk transaksi belanja online lazada khusus mode transfer bni dan bank lainnya tapi sekarang tidak bisa lagi tambah susah tolong solusinya dong .. bagus	[biasanya, bisa, dipakai, untuk, transaksi, belanja, online, lazada, khusus, mode, transfer, bni, dan, bank, lainnya, tapi, sekarang, tidak, bisa, lagi, tambah, susah, tolong, solusinya, dong]
	[bagus]

Hasil *stopwords removal* dapat dilihat pada Tabel VII.

TABEL VII
HASIL STOPWORDS REMOVAL

Sebelum Stopwords Removal	Setelah Stopwords Removal
[belum, pernah]	[]
[saya, sdh, upgrade, tp, pas, di, buka, kbagus, gak, bisa]	[sdh, upgrade, tp, pas, buka, kbagus, gak]
[sangat, ribet, pertama, saya, kebank, untuk, pergantian, kartu, atm, karna, rusak, kata, pihak, bank, kartu, terbatas, hanya, untuk, yg, sudah, tua, yg, usia, muda, direkomendasikan, menggunakan, brimo, trus, langsung, daftar, dics, karna, hp, gagal, ga, langsung, bilang, pas, libur, kerja, aja,	[ribet, kebank, pergantian, kartu, atm, karna, rusak, bank, kartu, terbatas, yg, tua, yg, usia, muda, direkomendasikan, brimo, trus, langsung, daftar, dics, karna, hp, gagal, ga, langsung, bilang, pas, libur, kerja, aja,

langsung, daftar, dics, karna, dari, hp, gagal, terus, dan, ga, langsung, jadi, katanya, harus, datang, lagi, lain, waktu, saya, bilang, saya, ada, waktu, pas, libur, kerja, hari, ini, aja, dan, nunggu, lagi, libur, mnggu, kemudian, dan, tetap, ga, bisa, hari, itu, juga, skrng, diulang, ulang, daftar, via, hp, gagal, terus, kalo, bisa, ribet, knapa, harus, mudah]	nunggu, libur, mnggu, ga, skrng, diulang, ulang, daftar, via, hp, gagal, kalo, ribet, knapa, mudah]
[biasanya, bisa, dipakai, untuk, transaksi, belanja, online, lazada, khusus, mode, transfer, bni, dan, bank, lainnya, tapi, sekarang, tidak, bisa, lagi, tambah, susah, tolong, solusinya, dong]	[dipakai, transaksi, belanja, online, lazada, khusus, mode, transfer, bni, bank, susah, tolong, solusinya]
[bagus]	[bagus]

Hasil *stemming* dapat dilihat pada Tabel VIII.

TABEL VIII
HASIL STEMMING

Sebelum Stemming	Setelah Stemming
[]	[]
[sdh, upgrade, tp, pas, buka, kbagus, gak]	[sdh, upgrade, tp, pas, buka, kbagus, gak]
[ribet, kebank, pergantian, kartu, atm, karna, rusak, bank, kartu, terbatas, yg, tua, yg, usia, muda, direkomendasikan, brimo, trus, langsung, daftar, dics, karna, hp, gagal, ga, langsung, bilang, pas, libur, kerja, aja, nunggu, libur, mnggu, ga, skrng, diulang, ulang, daftar, via, hp, gagal, kalo, ribet, knapa, mudah]	[ribet, bank, ganti, kartu, atm, karna, rusak, bank, kartu, batas, yg, tua, yg, usia, muda, rekomendasi, brimo, trus, langsung, daftar, dics, karna, hp, gagal, ga, langsung, bilang, pas, libur, kerja, aja, nunggu, libur, mnggu, ga, skrng, ulang, ulang, daftar, via, hp, gagal, kalo, ribet, knapa, mudah]
[dipakai, transaksi, belanja, online, lazada, khusus, mode, transfer, bni, bank, susah, tolong, solusinya]	[pakai, transaksi, belanja, online, lazada, khusus, mode, transfer, bni, bank, susah, tolong, solusi]
[bagus]	[bagus]

Setelah semua tahapan *pre-processing* tersebut dilakukan, terdapat 204 baris kosong di kolom *content*. Peneliti kemudian memfilter baris-baris kosong tersebut, dan didapatkan hasil akhir jumlah data adalah 9138 baris, dengan rincian 2997 untuk label positif, 3041 untuk label netral, dan 3100 untuk label negatif. Hasil *filtering* baris kosong dapat dilihat pada Tabel IX.

TABEL IX
HASIL FILTERING BARIS KOSONG

Sebelum difilter	Setelah difilter
[]	[sdh, upgrade, tp, pas, buka, kbagus, gak]
[sdh, upgrade, tp, pas, buka, kbagus, gak]	[ribet, bank, ganti, kartu, atm, karna, rusak, bank, kartu, batas, yg, tua, yg, usia, muda, rekomendasi, brimo, trus, langsung, daftar, dics, karna, hp, gagal, ga, langsung, bilang, pas, libur, kerja, aja, nunggu, libur, mnggu, ga, skrng, ulang, ulang, daftar, via, hp, gagal, kalo, ribet, knapa, mudah]
[ribet, bank, ganti, kartu, atm, karna, rusak, bank, kartu, batas, yg, tua, yg, usia, muda, rekomendasi, brimo, trus, langsung, daftar, dics, karna, hp, gagal, ga, langsung, bilang, pas, libur, kerja, aja, nunggu, libur, mnggu, ga, skrng, ulang, ulang, daftar, via, hp, gagal, kalo, ribet, knapa, mudah]	[pakai, transaksi, belanja, online, lazada, khusus, mode, transfer, bni, bank, susah, tolong, solusi]
[bagus]	[aplikasi, bagus, pendaptaran, siang, gagal, hapus, download, ulang, ulang, coba, daptar, malam, hasil, mudah, cepat, perosesnya, mudah, mudah, aplikasi, aman, tingkat, keamanannya]

Selanjutnya adalah tahapan *mapping* (pemetaan). Tahapan ini memetakan kolom label menjadi nilai numerik. Label negatif bernilai 0, netral bernilai 1, dan positif bernilai 2. Hasil *pre-processing* dan *mapping* kemudian disimpan menjadi *Dataframe* final, yang dapat dilihat pada Tabel X.

TABEL X
HASIL MAPPING

content	score	label
[sdh, upgrade, tp, pas, buka, kbagus, gak]	3	1
[ribet, bank, ganti, kartu, atm, karna, rusak, bank, kartu, batas, yg, tua, yg, usia, muda, rekomendasi, brimo, trus, langsung, daftar, dics, karna, hp, gagal, ga, langsung, bilang, pas, libur, kerja, aja, nunggu, libur, mnggu, ga, skrng, ulang, ulang, daftar, via, hp, gagal, kalo, ribet, knapa, mudah]	1	0
[pakai, transaksi, belanja, online, lazada, khusus, mode, transfer, bni, bank, susah, tolong, solusi]	3	1
[bagus]	5	2

[aplikasi, bagus, pendaptaran, siang, gagal, hapus, downlod, ulang, ulang, coba, daptar, malam, hasil, mudah, cepat, perosesnya, mudah, mudah, aplikasi, aman, tingkat, keamananya]	5	2
---	---	---

Hasil *pre-processing* kemudian dibuatkan *word clouds*. *Word clouds* adalah representasi visual dari teks yang menampilkan kata-kata yang sering muncul dalam kumpulan data. *Word clouds* dibuat sebagai visualisasi yang mewakili setiap label yang ada [29]. Ukuran setiap kata dalam *word cloud* menunjukkan frekuensi atau pentingnya kata tersebut dalam teks.

Word cloud untuk sentimen positif dapat dilihat pada Gambar 5.



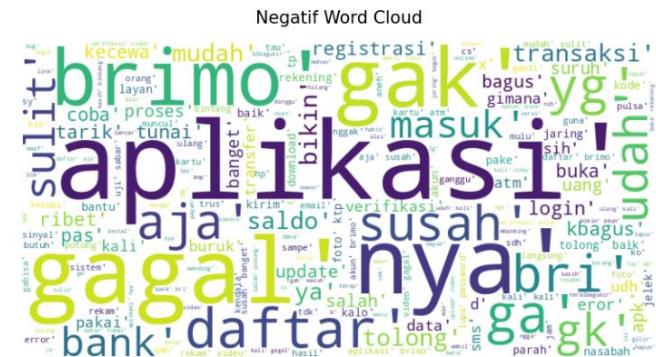
Gambar 5. Word Cloud Sentimen Positif

Word cloud untuk sentimen netral dapat dilihat pada Gambar 6.



Gambar 6. Word Cloud Sentimen Netral

Word cloud untuk sentimen negatif dapat dilihat pada Gambar 7.



Gambar 7. Word Cloud Sentimen Negatif

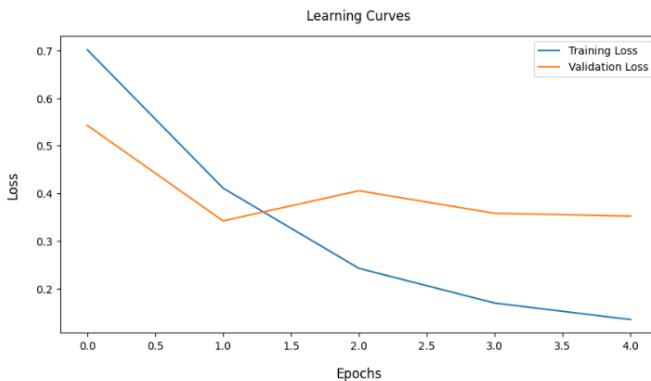
Setelah itu, data dari hasil *pre-processing* akan dibagi menjadi data latih (*training*), data validasi (*validation*), dan data uji (*test*), dengan rasio pembagian (*splitting*) adalah 80:10:10. Kemudian dilakukan proses tokenisasi menggunakan *IndoBERT Tokenizer* dengan *max length* sebesar 72, dan *data loaders* dengan *batch size* sebesar 8. Proses pelatihan dilakukan selama 5 *epoch*, serta memanfaatkan *optimizer* AdamW dengan *learning rate* 1e-5 dan *weight decay* sebesar 0.01. Pelatihan juga menerapkan pembelajaran bertahap (*learning rate scheduler*), dan mekanisme penghentian dini (*early stopping*) untuk mencegah *overfitting*. *Early stopping* telah dipelajari dengan baik untuk mengurangi *overfitting* dan meningkatkan generalisasi [30]. Selain itu, lapisan *dropout* juga ditambahkan untuk meningkatkan generalisasi model. Hasil pelatihan model dapat dilihat pada Tabel XI.

TABEL XI
HASIL PELATIHAN MODEL

Epoch	Train Loss	Validation Loss
1/5	0.7015	0.5428
2/5	0.4109	0.3422
3/5	0.2426	0.4055
4/5	0.1698	0.3582
5/5	0.1351	0.3523

Hasil pelatihan model menunjukkan bahwa pada *epoch* pertama, *train loss* adalah 0.7015 dan *validation loss* adalah 0.5428. Pada *epoch* kedua, *train loss* menurun menjadi 0.4109 dan *validation loss* menurun menjadi 0.3422, menunjukkan bahwa model belajar dengan baik dari data pelatihan. Namun, pada *epoch* ketiga, *validation loss* meningkat menjadi 0.4055 meskipun *train loss* terus menurun menjadi 0.2426. Hal ini mengindikasikan bahwa model mulai mengalami *overfitting*. Pada *epoch* keempat, *validation loss* kembali menurun menjadi 0.3582 sementara *train loss* menjadi 0.1698. Pada *epoch* kelima, *validation loss* sedikit menurun menjadi 0.3523 dengan *train loss* yang lebih rendah yaitu 0.1351. Dengan mekanisme *early stopping*, pelatihan dihentikan setelah *epoch* kelima karena tidak ada peningkatan signifikan pada *validation loss*, mengindikasikan bahwa

model telah mencapai titik optimalnya. Grafik *learning curves* untuk *training loss* dan *validation loss* dapat dilihat pada Gambar 8.



Gambar 8. *Learning Curves*

Gambar *learning curves* di atas menunjukkan tren penurunan yang konsisten dalam *training loss* selama pelatihan, yang menunjukkan bahwa model berhasil belajar dari data pelatihan. Namun, *validation loss* menunjukkan beberapa fluktuasi, terutama peningkatan pada *epoch* ketiga, yang menandakan adanya potensi *overfitting*. Pada *epoch* keempat dan kelima, *validation loss* kembali menurun, meskipun tidak secepat penurunan *train loss*. Perbedaan antara *train loss* dan *validation loss* yang stabil menunjukkan bahwa model cukup baik dalam generalisasi, meskipun ada indikasi awal dari *overfitting*.

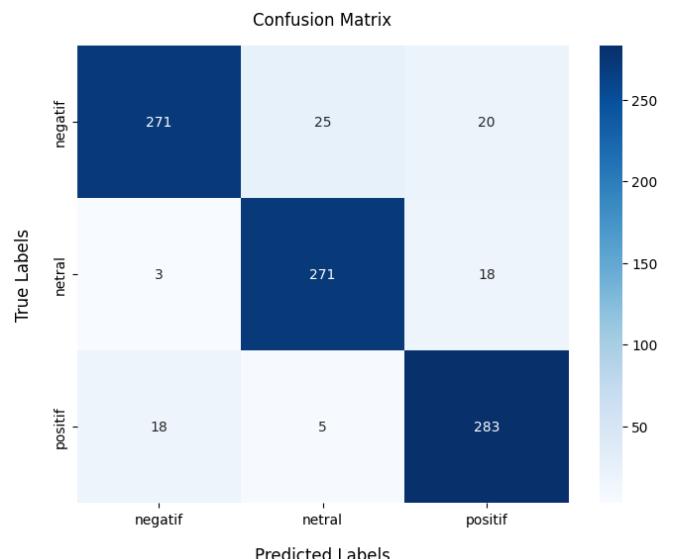
Setelah pelatihan selesai, model dievaluasi menggunakan data uji dan menghasilkan *classification report* berupa metrik-metrik evaluasi seperti *accuracy*, *precision*, *recall*, *f1-score*, dan *confusion matrix*. Berdasarkan hasil evaluasi, model menunjukkan kinerja yang baik dengan akurasi keseluruhan sebesar 90%. Rincian *classification report* dapat dilihat pada Tabel XII.

TABEL XII
CLASSIFICATION REPORT

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>
negatif	0.93	0.86	0.89
netral	0.90	0.93	0.91
positif	0.88	0.92	0.90
accuracy		0.90	

Evaluasi model menghasilkan *precision* 0.93 untuk kelas negatif, 0.90 untuk kelas netral, dan 0.88 untuk kelas positif. *Recall* yang dihasilkan adalah 0.86 untuk kelas negatif, 0.93 untuk kelas netral, dan 0.92 untuk kelas positif. Sedangkan hasil *f1-score* adalah 0.89 untuk kelas negatif, 0.91 untuk kelas netral, dan 0.90 untuk kelas positif. Metrik-metrik ini menunjukkan bahwa model memiliki keseimbangan yang baik antara *precision* dan *recall* untuk setiap kelas, yang mengarah pada *f1-score* yang cukup tinggi. Hasil tersebut juga menunjukkan bahwa *IndoBERT* memiliki performa yang lebih baik dibandingkan *Naïve Bayes* dan masih sebanding

dengan *SVM*, namun unggul dalam memahami konteks bahasa Indonesia secara lebih baik. *Confusion matrix* hasil evaluasi model dapat dilihat pada Gambar 9.



Gambar 9. *Confusion Matrix*

Gambar *confusion matrix* menunjukkan distribusi prediksi model terhadap label yang sebenarnya [31]. Dari *confusion matrix*, dapat dilihat bahwa model memiliki kesalahan prediksi yang relatif kecil. Misalnya, untuk kelas negatif, ada 271 prediksi yang benar, 25 prediksi salah ke kelas netral, dan 20 prediksi salah ke kelas positif. Untuk kelas netral, ada 271 prediksi yang benar, 3 prediksi salah ke kelas negatif, dan 18 prediksi salah ke kelas positif. Untuk kelas positif, ada 283 prediksi yang benar, 5 prediksi salah ke kelas netral, dan 18 prediksi salah ke kelas negatif.

Selain mengevaluasi performa *IndoBERT*, penelitian ini juga menganalisis sentimen pengguna terhadap aplikasi BRImo. Sentimen positif umumnya berisi apresiasi terhadap fitur transaksi yang cepat dan tampilan antarmuka yang *user-friendly*. Ulasan netral banyak berisi komentar tentang fitur tambahan yang diharapkan pengguna, seperti peningkatan fitur notifikasi transaksi dan kemudahan integrasi dengan dompet digital lainnya. Sementara itu, sentimen negatif didominasi oleh keluhan mengenai kegagalan *login*, transaksi yang tertunda atau gagal, serta layanan pelanggan yang lambat dalam merespons permasalahan.

Hasil analisis sentimen ini dapat menjadi bahan masukan bagi BRI dan dapat dimanfaatkan untuk meningkatkan layanan BRImo dengan berbagai cara. Misalnya, peningkataan stabilitas *server* dapat mengurangi kendala *login* dan transaksi yang tertunda. Analisis otomatis memungkinkan pengembang untuk lebih cepat mengidentifikasi ulasan negatif dan merespons keluhan pengguna secara lebih efektif. Sentimen positif dapat digunakan untuk mengidentifikasi fitur yang disukai pengguna, sementara sentimen netral dan negatif dapat memberikan wawasan mengenai aspek yang perlu diperbaiki.

Dengan menerapkan hasil analisis ini, BRImo dapat meningkatkan kepuasan pengguna serta merancang strategi pengembangan yang lebih responsif dan berbasis data.

IV. KESIMPULAN

Melalui serangkaian tahapan yang meliputi pengumpulan data, pra-pemrosesan, pelatihan model, dan evaluasi, penelitian ini berhasil mengembangkan model yang mampu mengklasifikasikan sentimen ulasan pengguna dengan akurasi yang cukup tinggi. Hasil akhir menunjukkan bahwa model *IndoBERT* yang diterapkan mampu memberikan akurasi keseluruhan sebesar 90%, dengan nilai *precision*, *recall*, dan *f1-score* yang seimbang untuk setiap kategori sentimen. Hasil tersebut menunjukkan bahwa model *IndoBERT* yang dilatih dengan dataset ulasan aplikasi BRImo mampu menghasilkan prediksi sentimen yang cukup akurat.

Keberhasilan model ini menunjukkan bahwa *IndoBERT* memiliki keunggulan dalam memahami konteks bahasa Indonesia, khususnya dalam ulasan-ulasan aplikasi perbankan digital yang sering kali menggunakan bahasa informal dan campuran. Hasil ini juga menggarisbawahi pentingnya proses pra-pemrosesan data yang tepat, termasuk *balancing* data, *cleaning*, *tokenizing*, dan penggunaan teknik-teknik seperti *stemming* dan *stopwords removal* untuk meningkatkan kualitas data yang digunakan dalam pelatihan model.

Meskipun terdapat indikasi awal dari *overfitting*, penggunaan mekanisme *early stopping* dan *dropout* membantu dalam mengurangi efek *overfitting* dan meningkatkan generalisasi model. Dengan tren penurunan yang konsisten dalam *training loss* dan *validation loss* yang stabil, model ini menunjukkan performa yang baik dalam analisis sentimen. Evaluasi lebih lanjut menggunakan metrik *precision*, *recall*, *f1-score*, dan *confusion matrix* memberikan gambaran yang lebih lengkap tentang kinerja model.

Selain evaluasi model, penelitian ini juga mengungkapkan bahwa sebagian besar ulasan negatif berkaitan dengan masalah teknis seperti kegagalan *login* dan transaksi, yang dapat menjadi fokus utama perbaikan bagi pengembang BRImo. Hasil analisis ini dapat diterapkan dalam peningkatan layanan BRImo melalui optimasi fitur, peningkatan respons terhadap ulasan negatif, dan penyempurnaan pengalaman pengguna secara keseluruhan.

Namun demikian, penelitian ini juga memiliki keterbatasan. *Dataset* yang digunakan dalam penelitian ini hanya mencakup ulasan dari *Google Play Store*, yang mungkin tidak mewakili semua persepsi pengguna aplikasi BRImo. Penelitian lebih lanjut dapat mengintegrasikan ulasan dari berbagai platform untuk mendapatkan gambaran yang lebih komprehensif. Penggunaan data tambahan dari berbagai sumber dapat membantu model dalam menangani variasi bahasa dan konteks yang lebih luas.

Secara keseluruhan, penelitian ini berhasil menunjukkan potensi *IndoBERT* dalam analisis sentimen untuk ulasan aplikasi perbankan digital, serta memberikan wawasan penting bagi pengembang aplikasi untuk meningkatkan layanan berdasarkan umpan balik pengguna.

Hasil penelitian ini diharapkan dapat menjadi pijakan untuk penelitian-penelitian selanjutnya dalam bidang analisis sentimen berbasis bahasa Indonesia dan aplikasi perbankan digital. Ke depannya, peningkatan lebih lanjut dapat dilakukan dengan mencoba teknik *balancing* data yang lebih baik seperti *SMOTE* untuk *oversampling*, peningkatan *pre-processing*, atau menggunakan arsitektur model yang lebih kompleks. Selain itu, studi mendatang dapat lebih mendalamai aspek spesifik dalam sentimen negatif guna memberikan rekomendasi yang lebih tepat bagi pengembang aplikasi perbankan digital.

DAFTAR PUSTAKA

- [1] P. A. T. P. Pika, K. Darmaastawan, A. I. Dewiningrat, and J. J. P. Latupeirissa, “Antesenden Minat Menggunakan Bri Mobile (Brimo) Pada Nasabah Bri Kantor Cabang Sesetan, Denpasar,” *Jurnal Akuntansi dan Pajak*, vol. 22, no. 2, 2022, doi: <http://dx.doi.org/10.29040/jap.v22i2.4439>.
- [2] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, “Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest,” 2022. [Online]. Available: <http://j-ptik.ub.ac.id>
- [3] S. A. Prasetyo and W. T. Atmojo, “Jurnal Media Informatika Budidarma Analisis Sentimen Terhadap Penggunaan Aplikasi Astro Menggunakan Algoritma Naïve Bayes,” vol. 7, pp. 2149–2158, 2023, doi: 10.30865/mib.v7i4.6750.
- [4] S. M. Kom, “Sentimen Pengguna Aplikasi BRImo: Kinerja Algoritma Support Vector Machine, Naive Bayes, dan Adaboost,” *SATIN - Sains dan Teknologi Informasi*, vol. 9, no. 2, pp. 195–204, Dec. 2023, doi: 10.33372/stn.v9i2.1057.
- [5] N. Fibriyanti Arminda, N. Sulistiyati, and T. Nur Padilah, “Implementasi Algoritma Multinomial Naive Bayes Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo,” 2023.
- [6] M. Khoirul, U. Hayati, and O. Nurdianwan, “Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naïve Bayes,” 2023.
- [7] M. Umair and E. R. Sutanto, “Jurnal Media Informatika Budidarma Analisis Sentimen Ulasan Pengguna Pada Aplikasi BRImo BRI Menggunakan Metode Klasifikasi Algoritma Naïve Bayes,” 2024, doi: 10.30865/mib.v8i2.7381.
- [8] A. Nuraini, A. Faqih, G. Dwilestari, N. D. Nuris, and R. Narasati, “Analisis Sentimen Terhadap Review Aplikasi Brimo Di Google Play Store Menggunakan Algoritma Naïve Bayes,” 2023.
- [9] P. S. Arya Bhagaskara M, S. Suryani Prasetyowati, and Y. Sibaroni, “Hoax Detection of Indonesian News Media on Twitter Using IndoBERT with Word Embedding Word2Vec,” vol. 7, no. 3, pp. 1088–1096, 2023, doi: 10.30865/mib.v7i3.6367.
- [10] R. Nugroho, N. Azka, and W. Sayudha Rendra Graha, “Analisis Sentimen Ulasan Aplikasi Mobile JKN di Google PlayStore Menggunakan IndoBERT.”
- [11] G. Hakim, T. N. Fatyanosa, and A. W. Widodo, “Analisis Sentimen Masyarakat terhadap Kereta Cepat Whoosh pada Platform X menggunakan IndoBERT,” 2017. [Online]. Available: <http://j-ptik.ub.ac.id>
- [12] D. Sandi and E. Utami, “Analisis Sentimen Publik Terhadap Elektabilitas Ganjar Pranowo di Tahun Politik 2024 di Twitter dengan Algoritma KNN dan Naïve Bayes,” vol. 7, no. 3, pp. 1097–1108, 2023, doi: 10.30865/mib.v7i3.6298.
- [13] C. Solheim Bojer and J. P. Meldgaard, “Learnings from Kaggle’s Forecasting Competitions.” [Online]. Available: <https://www.kaggle.com/>
- [14] H. Parasian Doloksaribu and Y. T. Samuel, “Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi,” vol. 16, no. 1, 2022, doi: 10.47111/JTI.
- [15] J. Prasetya, “Leibniz : Jurnal Matematika Penerapan Klasifikasi Naïve Bayes Dengan Algoritma Random Oversampling Dan

- [16] Random Undersampling Pada Data Tidak Seimbang Cervical Cancer Risk Factors,” vol. 2, no. 2.
- [17] R. Merdiansah and A. Ali Ridha, “Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT,” *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, vol. 7, no. 1, pp. 221–228, 2024.
- [18] R. Nurul Ikhnsani and F. Fauzi Abdulloh, “Jurnal Media Informatika Budidarma Optimasi SVM dan Decision Tree Menggunakan SMOTE Untuk Mengklasifikasi Sentimen Masyarakat Mengenai Pinjaman Online,” vol. 7, pp. 1667–1677, 2023, doi: 10.30865/mib.v7i4.6809.
- [19] F. Putra Herlambang and D. Avianto, “Jurnal Media Informatika Budidarma Analisis Sentimen Opini Pengguna Twitter Terhadap Tragedi Kanjuruhan Malang dengan Metode Support Vector Machine,” vol. 7, pp. 1727–1739, 2023, doi: 10.30865/mib.v7i4.6332.
- [20] L. Yusuf and S. Masripah, “Sentimen Analisis Chatgpt Dengan Algoritma Naïve Bayes Dan Optimasi PSO,” *INTI Nusa Mandiri*, vol. 18, no. 1, pp. 59–64, Aug. 2023, doi: 10.33480/inti.v18i1.4230.
- [21] T. T. Mengistie and D. Kumar, “Deep Learning Based Sentiment Analysis on COVID-19 Public Reviews,” in *3rd International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2021*, Institute of Electrical and Electronics Engineers Inc., Apr. 2021, pp. 444–449. doi: 10.1109/ICAIIC51459.2021.9415191.
- [22] R. Maulana, A. Voutama, and T. Ridwan, “Analisis Sentimen Ulasan Aplikasi Mypertamina Pada Google Play Store Menggunakan Algoritma NBC,” *Jurnal Teknologi Terpadu*, vol. 9, no. 1, pp. 42–48, 2023.
- [23] A. Diffa Maulana and K. M. Lhaksmana, “Jurnal Media Informatika Budidarma Sentiment Analysis on Tweets of Kanjuruhan Tragedy Using Deep Learning IndoBERTweet,” 2023, doi: 10.30865/mib.v7i3.6115.
- [24] D. Ismiyana Putri, A. Nurul Alfian, M. Yudhi Putra, and P. Dwi Mulyo, “IndoBERT Model Analysis: Twitter Sentiments on Indonesia’s 2024 Presidential Election,” 2024. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [25] P. Hanif Zakaria, D. Nurjannah, and H. Nurrahmi, “Misogyny Text Detection on Tiktok Social Media in Indonesian Using the Pre-trained Language Model IndoBERTweet,” vol. 7, no. 3, pp. 1297–1305, 2023, doi: 10.30865/mib.v7i3.6438.
- [26] L. Geni, E. Yulianti, and D. I. Sensuse, “Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models,” *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, vol. 9, no. 3, pp. 746–757, 2023, doi: 10.26555/jiteki.v9i3.26490.
- [27] N. B. B. Tjikdaphia and S. Sulastri, “Comparison Of Nbc, Svm, Knn Classification Results In Sentiment Analysis Of Mobile JKN,” *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 9, no. 4, pp. 665–672, Sep. 2023, doi: 10.33330/jurteksi.v9i4.2539.
- [28] D. T. Dunrui, “brimo_googleplaystore_review,” 2023. Accessed: Jul. 11, 2024. [Online]. Available: https://www.kaggle.com/datasets/dannytheodore/brimo-app-review?resource=download&select=brimo_googleplaystore_review.csv
- [29] F. Novianti and K. R. N. Wardani, “Analisis Sentimen Masyarakat Terhadap Data Tweet Traveloka Selama Rapid Test Antigen Menggunakan Algoritma Naïve Bayes,” *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 3, pp. 922–933, Aug. 2023, doi: 10.29100/jipi.v8i3.3973.
- [30] L. Rice, E. Wong, and J. Z. Kolter, “Overfitting in adversarially robust deep learning,” 2020. [Online]. Available: <https://github.com/>
- [31] R. Hardian, L. D. Oktaviana, and A. Hamdi, “Sentiment Analysis Of Pegipegi.Com On Google Playstore With Naïve Bayes Algorithm,” *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 10, no. 3, pp. 583–590, Jun. 2024, doi: 10.33330/jurteksi.v10i3.3201.