

ANALISIS SENTIMEN ULASAN PENGGUNA GOJEK MENGUNAKAN ALGORITMA NAÏVE BAYES DAN BIGRU

Korin Kusuma Ningrum, Agung Kharisma Hidayah

Sistem Informasi, Universitas Muhammadiyah Bengkulu

Jalan Bali, Kampung Bali, Teluk Segara, Bengkulu, 38119, Indonesia

kusumakorin@gmail.com

ABSTRAK

Transportasi ojek online merupakan salah satu moda transportasi yang banyak digunakan di Indonesia. Gojek sebagai salah satu penyedia layanan utama menghadapi persaingan dari platform sejenis seperti Maxim dan Grab sehingga pemahaman terhadap persepsi pengguna aplikasi menjadi penting dalam upaya peningkatan kualitas layanan. Persepsi tersebut dapat diperoleh melalui ulasan pengguna yang tersedia pada Google Play Store dan merefleksikan pengalaman nyata selama menggunakan aplikasi. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Gojek dengan mengklasifikasikan sentimen positif dan negatif menggunakan algoritma *Naïve Bayes* dan *Bidirectional Gated Recurrent Unit* (BiGRU). Data penelitian diperoleh melalui teknik *web scraping* sebanyak 5.000 ulasan pengguna, kemudian diproses melalui tahap *preprocessing* yang meliputi *cleaning*, *case folding*, *tokenization*, *stopword removal*, dan *stemming*. Representasi data dilakukan menggunakan TF-IDF pada model *Naïve Bayes*, serta tokenisasi, *padding*, dan *embedding* pada model BiGRU sebelum diterapkan dalam proses klasifikasi. Evaluasi kinerja model menggunakan *confusion matrix* menunjukkan bahwa *Naïve Bayes* mencapai akurasi sebesar 92%, sedangkan BiGRU memperoleh akurasi sebesar 90%. Hasil penelitian menunjukkan bahwa *Naïve Bayes* memiliki efisiensi dan konsistensi klasifikasi yang lebih baik, sementara BiGRU tetap mampu menangkap pola urutan teks secara efektif, sehingga temuan ini dapat dijadikan dasar dalam pengambilan keputusan strategis untuk meningkatkan mutu pelayanan Gojek secara sistematis dan terukur.

Kata kunci : Analisis Sentimen; Gojek; *Naïve Bayes*; Ulasan Pengguna; BiGRU

1. PENDAHULUAN

Transportasi ialah bagian dari keperluan pokok masyarakat modern guna menunjang mobilitas sehari-hari. Di Indonesia, ojek menjadi moda transportasi yang populer karena kecepatan, fleksibilitas, serta kemampuannya menembus kemacetan perkotaan. Ojek awalnya beroperasi di pangkalan dengan tarif bervariasi dan pembayaran langsung. Perkembangan teknologi mendorong transformasi ojek konvensional menjadi ojek online, yaitu layanan yang memanfaatkan aplikasi berbasis smartphone yang dapat dipesan melalui aplikasi, dengan tarif jelas dan pembayaran fleksibel, sehingga lebih praktis dan aman [1].

Gojek ialah bagian dari aplikasi ojek online paling populer di Indonesia serta menawarkan sejumlah layanan seperti transportasi, pengiriman makanan, pembayaran digital, serta sejumlah layanan lain [2]. Keunggulan Gojek terletak pada kemudahan penggunaan, kecepatan layanan, serta inovasi yang berkelanjutan. Namun, keberadaan aplikasi sejenis seperti Maxim, Grab, dan layanan lokal lainnya meningkatkan tingkat persaingan secara signifikan. Hal ini menekankan pentingnya memahami sentimen pengguna terhadap kinerja Gojek agar data ulasan yang akurat dapat dimanfaatkan guna membantu peningkatan mutu pelayanan.

Pengguna dapat memberikan ulasan serta penilaian terhadap aplikasi melalui Google Play Store, sebuah platform resmi untuk pengunduhan, pemasangan, dan evaluasi aplikasi berbasis Android

[3]. Penelitian ini menggunakan data ulasan terbaru aplikasi Gojek, dengan jumlah unduhan lebih dari 100 juta dan rating ulasan rata-rata sebesar 4,7, yang memperlihatkan tingkat kepuasan pengguna secara umum. Sebagian besar ulasan berbentuk teks panjang, tidak terstruktur, dan jumlahnya sangat besar, sehingga pengolahan secara manual menjadi tidak efektif. Oleh karena itu, diperlukan analisis sentimen, yakni teknik guna menemukan serta mengelompokkan opini berbentuk teks berdasarkan sikap positif, negatif atau netral terhadap suatu topik [4]. Dengan penerapan metode ini, informasi penting dari ulasan dapat diperoleh secara cepat, akurat, dan sistematis, sehingga pengembang Gojek dapat memahami persepsi pengguna dengan keseluruhan serta serta pengambilan tindakan pembenahan atas layanan yang tepat.

Pada pengolahan ulasan berskala besar, dibutuhkan algoritma pembelajaran mesin yang mampu menangani data teks tidak terstruktur secara efektif. Penelitian ini menggunakan algoritma *Naïve Bayes* dan *Bidirectional Gated Recurrent Unit* (BiGRU) sebagai metode utama. *Naïve Bayes* dipilih karena efisiensi komputasinya serta kemampuannya belajar secara efektif dari data teks yang terbatas. Sedangkan BiGRU, yang merupakan jenis *Recurrent Neural Network* (RNN), dibuat guna pelaksanaan proses data secara berurutan melalui kemampuan menangkap pola kompleks dari kedua arah, yaitu maju dan mundur [5].

Perolehan data berupa ulasan pengguna aplikasi Gojek dilaksanakan terlebih dahulu dari Google Play Store dengan memanfaatkan teknik *web scraping*, sebelum keseluruhan proses tersebut diposisikan sebagai tahapan awal studi. Data yang terkumpul selanjutnya melewati proses *preprocessing* guna meningkatkan mutu teks, yang mencakup *cleaning*, *case folding*, *tokenization*, *stopword removal*, dan *stemming*. Selanjutnya, representasi data teks dibentuk ke dalam vektor dengan memanfaatkan metode TF-IDF sebagai masukan bagi model *Naïve Bayes*. Sedangkan bagi BiGRU, setelah tahap *preprocessing* dilakukan tokenisasi, *padding*, dan *embedding* agar data dapat diproses oleh jaringan saraf. Setelah tahap representasi data, klasifikasi sentimen dilakukan menggunakan kedua algoritma, dan kinerja pemodelan dievaluasi melalui *confusion matrix* guna melaksanakan pengukuran akurasi serta performa klasifikasi.

Studi ini mempunyai tujuan guna menganalisa sentimen ulasan pengguna aplikasi Gojek di Google Play Store menggunakan algoritma *Naïve Bayes* dan BiGRU dalam mengklasifikasikan ulasan menjadi sentimen positif dan negatif, sekaligus membandingkan kinerja kedua algoritma dalam hal akurasi dan performa klasifikasi. Hasil penelitian diharapkan memberikan informasi yang akurat terkait anggapan pengguna bagi layanan Gojek, yang nantinya bisa dapat menjadi landasan atas diambilnya keputusan strategis untuk meningkatkan kualitas layanan.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Sejumlah studi sebelumnya telah mengkaji analisis sentimen melalui pemanfaatan pendekatan pembelajaran mesin dan *deep learning*, termasuk penggunaan algoritma *Naïve Bayes* dan *Bidirectional Gated Recurrent Unit* (BiGRU). Salah satu penelitian menganalisis sentimen atas ulasan pengguna dari aplikasi Gojek di Play Store yang memanfaatkan algoritma *Naïve Bayes*. Pengujian pada 800 ulasan menghasilkan akurasi 68%, yang memperlihatkan bahwasanya cara ini cukup efektif guna mengklasifikasikan sentimen positif serta negatif terkait layanan Gojek [6].

Penelitian berikutnya, menganalisis sentimen ulasan pengguna aplikasi Gojek di Google Play Store melalui perbandingan *Naïve Bayes*, *Decision Tree*, serta KNN. Hasilnya memperlihatkan *Naïve Bayes* memiliki akurasi tertinggi (89%), sehingga lebih efektif dibandingkan metode lainnya [7]. Pendekatan *deep learning* juga diterapkan dalam analisis sentimen cuitan Twitter terkait COVID-19 dengan menggunakan model BiLSTM dan BiGRU berbasis GloVe Twitter. Hasilnya memperlihatkan BiGRU dua lapis memiliki akurasi validasi lebih tinggi dan stabil (86,78%) dibanding BiLSTM (84,91%), sehingga lebih efektif pada data media sosial [8].

Pada beberapa penelitian sebelumnya, teknik pembelajaran mesin memperlihatkan kinerja yang lebih baik dibandingkan pendekatan *deep learning* dalam analisis sentimen dan klasifikasi teks. Salah satunya ditunjukkan pada analisis sentimen komentar mahasiswa terhadap layanan daring FTK Universitas Pendidikan Ganesha selama pandemi COVID-19 dengan membandingkan *Naïve Bayes* dan LSTM. Pengujian terhadap 425 komentar memperlihatkan bahwasanya *Naïve Bayes* mencapai ketepatan sebanyak 83,69%, lebih besar daripada LSTM yang hanya mencapai 53,12%, sehingga lebih efektif pada dataset tersebut [9]. Temuan serupa juga terlihat pada klasifikasi berita hoaks, di mana perbandingan antara SVM, *Multinomial Naïve Bayes*, dan LSTM pada 2.000 data hoaks memperlihatkan bahwasanya SVM memiliki kinerja terbaik dengan f1-score sebesar 0,74, diikuti *Multinomial Naïve Bayes* sebesar 0,62, sedangkan LSTM memperlihatkan kinerja terendah dengan akurasi 0,49 [10].

Berdasarkan penelitian sebelumnya, meskipun banyak studi telah mengkaji analisis sentimen menggunakan *Naïve Bayes* maupun BiGRU, belum ada penelitian yang secara spesifik membandingkan kedua metode tersebut pada ulasan aplikasi Gojek. Maka dari itu, studi ini mempunyai tujuan guna menganalisis sentimen ulasan pengguna Gojek di Google Play Store menggunakan *Naïve Bayes* dan BiGRU, dengan fokus mengklasifikasikan ulasan menjadi sentimen positif dan negatif, sekaligus membandingkan kinerja kedua algoritma dalam hal akurasi dan performa. Perolehan atas analisis diharapkan bisa menggambarkan dengan lebih komprehensif terkait opini pengguna serta dapat dijadikan acuan strategis bagi pengembang dalam meningkatkan kualitas layanan Gojek.

2.2. Analisis Sentimen

Analisis sentimen merupakan salah satu teknik dalam pengolahan teks yang digunakan untuk mengidentifikasi dan mengklasifikasikan kecenderungan emosi atau opini yang terkandung dalam data tekstual [11]. Teknik ini bekerja dengan menganalisis kata, frasa, maupun pola bahasa yang muncul dalam suatu teks untuk menentukan sikap pengguna terhadap suatu objek, layanan, atau peristiwa. Dalam praktiknya, analisis sentimen umumnya mengelompokkan data ke dalam kategori sentimen positif, negatif dan netral. Pada penelitian berbasis layanan digital, analisis sentimen dimanfaatkan untuk memahami persepsi pengguna secara sistematis, sehingga hasil analisis dapat dijadikan dasar evaluasi dan peningkatan kualitas layanan [12].

2.3. Google Play Store

Ulasan pengguna merupakan bentuk umpan balik yang diberikan setelah pengguna menggunakan suatu aplikasi atau layanan. Ulasan ini berisi pengalaman, tingkat kepuasan, serta keluhan yang disampaikan

secara langsung oleh pengguna [13]. Google Play Store menyediakan fitur ulasan dan penilaian yang memungkinkan pengguna menyampaikan pendapatnya secara terbuka dan beragam. Oleh karena itu, ulasan pengguna pada Google Play Store menjadi sumber data yang relevan dan representatif untuk analisis sentimen, karena mencerminkan persepsi nyata pengguna terhadap kualitas dan kinerja suatu aplikasi [14].

2.4. Naïve Bayes

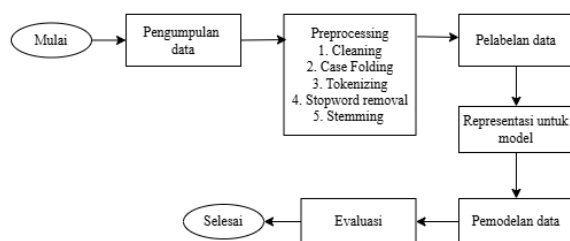
Naïve Bayes adalah algoritma klasifikasi *machine learning* yang diterapkan pada analisis sentimen untuk mengelompokkan teks berdasarkan probabilitas statistik suatu data. Algoritma ini bekerja dengan menghitung peluang kemunculan setiap kata terhadap masing-masing kelas sentimen berdasarkan data pelatihan. Hasil klasifikasi ditentukan oleh kelas dengan nilai probabilitas tertinggi. Meskipun memiliki asumsi independensi antar fitur, *Naïve Bayes* dikenal memiliki tingkat efisiensi yang tinggi dan performa yang stabil, sehingga sering digunakan dalam analisis sentimen untuk mengklasifikasikan teks ke dalam kategori sentimen positif maupun negatif [15].

2.5. BiGRU

Bidirectional Gated Recurrent Unit (BiGRU) merupakan algoritma *deep learning* yang ialah bagian dari variasi RNN yang dikembangkan guna menangani data berurutan seperti teks. Berbeda dengan GRU biasa, BiGRU bekerja dengan memproses urutan data dalam dua arah, yaitu dari awal ke akhir dan dari akhir ke awal, sehingga mampu menangkap konteks kata secara lebih menyeluruh. Dengan menggabungkan informasi dari kedua arah tersebut, BiGRU dapat memahami hubungan antar kata dalam suatu kalimat dengan lebih baik. Kemampuan ini menjadikan BiGRU efektif dalam analisis sentimen, khususnya untuk mengenali pola sentimen yang bergantung pada urutan dan konteks teks [16].

3. METODE PENELITIAN

Studi ini menganalisis ulasan pengguna aplikasi Gojek di Google Play Store memanfaatkan algoritma *Naïve Bayes* serta BiGRU, dengan beriktu langkah studi ditampilkan oleh Gambar 1.



Gambar 1. Tahapan Penelitian

3.1. Pengumpulan Data

Teknik *web scraping* yang dioperasikan melalui platform Google Colab dimanfaatkan pada studi ini guna mengumpulkan data. Data yang dikumpulkan mencakup 5.000 ulasan terbaru pengguna aplikasi Gojek yang berasal dari Google Play Store dan disimpan dalam format CSV. Ulasan tersebut merepresentasikan persepsi serta pengalaman pengguna terhadap layanan aplikasi Gojek dan digunakan sebagai data utama dalam proses analisis sentimen. Seluruh data yang telah dikumpulkan selanjutnya dipersiapkan untuk proses pengolahan dan analisis pada tahapan penelitian berikutnya.

3.2. Pre-Processing

Pasca tahapan dikumpulkannya data dilakukan, tahapan berikutnya dalam penelitian ini ialah *preprocessing* data teks ulasan untuk kedua model *Naïve Bayes* dan BiGRU. Tahapan mencakup *cleaning*, *case folding*, *tokenization*, *stopword removal*, serta *stemming*.

a. Cleaning

Tahap ini bertujuan penghapusan kriteria yang tidak selars, selayaknya tanda baca, angka, URL, emoji, serta simbol, yang menjadikan teks agar lebih bersih serta fokus pada informasi sentimen.

b. Case Folding

Case folding dilaksanakan dengan mengubah semua huruf hingga berbentuk huruf kecil guna menyamakan bentuk kata dan menghindari perbedaan akibat penggunaan huruf kapital.

c. Tokenization

Tokenization akan melaksanakan pemecahan teks ulasan agar berbentuk kata-kata (token) agar setiap kata dapat dianalisis dan digunakan sebagai fitur dalam proses selanjutnya.

d. Stopword Removal

Langkah ini melaksanakan penghilangan sejumlah kata umum yang acapkali terlihat, tapi mempunyai pengaruh kecil terhadap penentuan sentimen, misalnya 'dan', 'yang', 'di', atau 'ke'. Proses ini bertujuan menekan gangguan data (*noise*) dan meningkatkan akurasi analisis.

e. Stemming

Stemming mentransnformasikan kata dengan imbuhan hingga berbentuk kata dasar guna menggabungkan variasi kata dengan makna yang serupa serta meningkatkan akurasi analisis.

3.3. Pelabelan Data

Selanjutnya, dilakukan tahap pelabelan data dengan tiap-tiap ulasan dikelompokkan menjadi dua kelas sentimen, yakni positif serta negatif. Ulasan yang mempunyai rating 4 atau 5 diberikan label biner 1 untuk kelas positif, sedangkan ulasan dengan rating 1 hingga 3 diberikan label biner 0 untuk kelas negatif. Proses ini menggunakan pendekatan *rule-based* sehingga setiap ulasan memperoleh label secara otomatis, yang kemudian digunakan sebagai target dalam analisis sentimen.

3.4. Representasi untuk Model

Meskipun kedua algoritma melewati tahap *preprocessing* awal yang sama, *Naïve Bayes* menggunakan representasi TF-IDF, sedangkan BiGRU memerlukan tokenisasi numerik, *padding*, dan *embedding* untuk menangkap urutan kata dan konteks kalimat.

Pada model *Naïve Bayes*, setelah melalui tahap *preprocessing*, data teks ulasan selanjutnya dilaksanakan *scoring* melalui TF-IDF guna menggambarkan kata-kata penting berbentuk numerik yang siap digunakan pada proses klasifikasi sentimen.

$$W_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times \log\left(\frac{N}{df_t}\right) \quad (1)$$

Keterangan :

$W_{t,d}$: Bobot TF-IDF
 $tf_{t,d}$: Frekuensi kata pada dokumen
 idf_t : Inverse frekuensi dokumen
 df_t : Jumlah dokumen yang terkandung kata
 N : Jumlah keseluruhan dokumen

Sedangkan pada model BiGRU, setelah tahap *preprocessing* tersebut dilakukan tokenisasi untuk mengubah kata menjadi token numerik agar dapat diproses oleh jaringan saraf, *padding* menyesuaikan panjang setiap ulasan agar seragam, serta *embedding* mengubah token numerik menjadi vektor berdimensi tinggi yang menangkap konteks kata dalam ulasan.

3.5. Pemodelan Data

Pengelompokan sentimen pada data ulasan dilakukan melalui proses klasifikasi yang diarahkan pada dua kategori, yakni positif serta negatif, di mana mekanisme pemodelan ini dipahami terlebih dahulu sebelum diterapkan menggunakan algoritma *Naïve Bayes* serta BiGRU pada studi ini.

Naïve Bayes ialah algoritma pengkategorian probabilistik yang menerapkan *Teorema Bayes* melalui asumsi independensi tiap fitur. Cara ini mengkalkulasikan probabilitas setiap ulasan untuk masuk ke dalam kelas sentimen tertentu, kemudian menetapkan kelas melalui skor probabilitas tertinggi menjadi perolehan klasifikasi [17]. Berikut bentuk umum teorema *Naïve Bayes* dinyatakan:

$$P(C | X) = \frac{P(X | C) P(C)}{P(X)} \quad (2)$$

Keterangan :

C : Kelas sentimen target (positif atau negatif)
 X : Fitur teks dari ulasan
 $P(C | X)$: Probabilitas posterior (kelas yang dicari)
 $P(X | C)$: Probabilitas fitur X muncul di kelas C
 $P(C)$: Probabilitas awal kelas C
 $P(X)$: Probabilitas fitur X

BiGRU adalah model yang menggunakan GRU dua arah untuk memproses data berurutan. Setiap GRU memiliki dua gerbang utama, yaitu *reset gate* dan *update gate*, serta *candidate hidden state* yang

berperan dalam perhitungan *hidden state* pada setiap langkah urutan. Untuk mengurangi risiko *overfitting*, BiGRU menerapkan beberapa strategi, antara lain *Dropout*, regularisasi L2, *EarlyStopping*, *validation split*, serta pembatasan jumlah epoch selama pelatihan.

3.6. Evaluasi

Dalam langkah selanjutnya, kinerja pemodelan pada studi ini dievaluasi memanfaatkan *confusion matrix*, yang berfungsi untuk membandingkan hasil prediksi dengan label aktual. Penilaian kinerja model dilakukan melalui metrik *recall*, *precision*, *accuracy*, serta F1-score sebagaimana dirumuskan pada persamaan di bawah.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

Keterangan :

TP : Positif benar (*True Positif*)
 TN : Negatif benar (*True Negative*)
 FP : Negatif salah (*False Positive*)
 FN : Positif salah (*False Negative*)

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Data penelitian didapatkan lewat teknik *web scraping* menggunakan library Python di Google Colab. Pengumpulan data berupa ulasan pengguna aplikasi Gojek di Google Play Store dengan skala rating 1–5. Skala ini merepresentasikan pengalaman positif maupun negatif pengguna. Proses pengumpulan dilakukan untuk memperoleh sampel yang terkini dan representatif. Dari proses ini, terkumpul sebanyak 5.000 ulasan terbaru. Data kemudian dilaksanakan penyimpanan melalui format CSV guna memfasilitasi proses *preprocessing* dan analisis sentimen.

4.2. Preprocessing

Langkah selanjutnya adalah *preprocessing* guna mempersiapkan serta membersihkan data teks supaya siap dianalisis untuk model *Naïve Bayes* dan BiGRU. Proses ini mencakup *cleaning*, *case folding*, *tokenization*, *stopword removal*, serta *stemming*. Sebagaimana ditampilkan oleh Tabel 1.

Tabel 1. Tahap *Preprocessing*

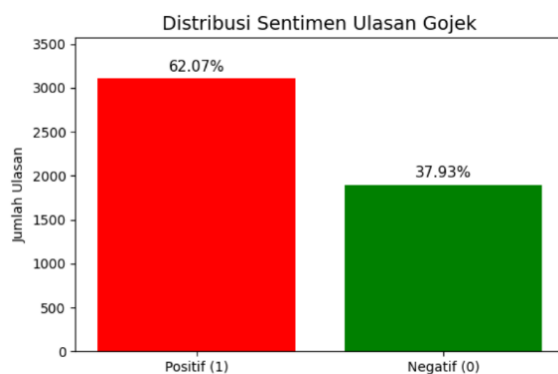
Tahapan	Hasil
Teks ulasan	Kalau Lagi hujan cari drivernya susah, sering lag, maps gapernah pas di titik alamatnya
Cleaning	Kalau Lagi hujan cari drivernya susah, sering lag, maps gapernah pas di titik alamatnya

Tahapan	Hasil
Case Folding	kalau lagi hujan cari drivernya susah, sering lag, maps gapernah pas di titik alamatnya
Tokenization	['kalau', 'lagi', 'hujan', 'cari', 'drivernya', 'susah', 'sering', 'lag', 'maps', 'gapernah', 'pas', 'di', 'titik', 'alamatnya']
Stopword Removal	['hujan', 'cari', 'drivernya', 'susah', 'lag', 'maps', 'gapernah', 'pas', 'titik', 'alamatnya']
Stemming	['hujan', 'cari', 'drivernya', 'susah', 'lag', 'map', 'gapernah', 'pa', 'titik', 'alamatnya']

Berdasarkan Tabel 1, teks ulasan diproses melalui *cleaning* untuk menghapus elemen yang tidak diperlukan, *case folding* berubah berbentuk huruf kecil, *tokenization* memisahkan kata, *stopword removal* menyaring kata kurang bermakna, dan *stemming* menghasilkan kata dasar, sehingga data siap dianalisis.

4.3. Pelabelan data

Berdasarkan distribusi hasil pelabelan, dari total 5.000 ulasan pengguna Gojek, sebanyak 3.104 ulasan (62.07%) termasuk dalam kelas positif, sedangkan 1.896 ulasan (37.93%) termasuk dalam kelas negatif. Hasil ini memperlihatkan bahwasanya sebagian besar ulasan bersifat positif, dan distribusi dataset yang cukup seimbang membuat data layak dimanfaatkan guna melatih serta menguji pemodelan *Naïve Bayes* maupun BiGRU, sebagaimana yang ditampilkan oleh Gambar 2.



Gambar 2. Hasil Pelabelan Data

Dari gambar 2 tersebut, memperlihatkan distribusi klasifikasi sentimen ulasan pengguna aplikasi Gojek. Dari total ulasan, 62.07% tergolong positif, sedangkan 37.93% bersifat negatif. Visualisasi ini menggambarkan persepsi pengguna secara keseluruhan, dengan mayoritas ulasan memperlihatkan sentimen positif terhadap layanan Gojek.

4.4. Representasi untuk Model

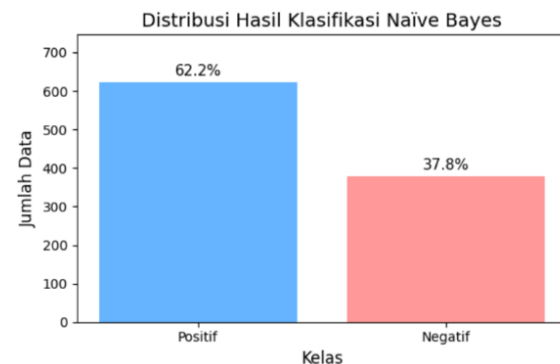
Tahap pembobotan data pada *Naïve Bayes* memanfaatkan metode TF-IDF, yakni dengan

menghitung frekuensi munculnya kata pada dokumen serta tingkat keunikannya pada seluruh kumpulan dokumen. Perkalian nilai TF dan IDF menghasilkan bobot kata yang merepresentasikan tingkat kepentingannya, sehingga mendukung proses analisis dan klasifikasi teks secara lebih efektif.

Pada pemodelan BiGRU, teks ulasan terlebih dahulu ditransformasikan menjadi bentuk numerik melalui proses tokenisasi, kemudian diseragamkan panjang urutannya menggunakan *padding*. Selanjutnya, data numerik tersebut dipetakan ke dalam vektor berdimensi 100 melalui *embedding layer* yang bersifat *trainable*, sehingga model dapat menangkap konteks kata sesuai dengan karakteristik ulasan dan meningkatkan kinerja klasifikasi sentimen.

4.5. Pemodelan Data

Dataset yang digunakan berjumlah 5.000 ulasan pengguna dilaksanakan pembagian menjadi 80% bagi pelatihan serta 20% bagi pengujian agar pemodelan dapat mempelajari pola sentimen dan dievaluasi performanya pada data baru. Pembagian ini diterapkan pada kedua algoritma, yakni *Naïve Bayes* dan BiGRU, untuk memastikan perbandingan kinerja yang konsisten.



Gambar 3. Hasil Pengujian *Naïve Bayes*

Mengacu pada Gambar 3, perolehan uji yang memanfaatkan algoritma *Naïve Bayes* memperlihatkan bahwasanya sebagian besar data uji diklasifikasikan ke dalam sentimen positif, yaitu sebesar 62.2%, sedangkan 37.8% data termasuk dalam sentimen negatif. Perolehan tersebut memperlihatkan bahwasanya model *Naïve Bayes* lebih banyak mengidentifikasi ulasan pengguna dengan sentimen positif.

Tabel 2. Hasil pengujian BiGRU

Model	Akurasi Training	Akurasi Validasi	Akurasi Uji	Loss Uji
BiGRU	90.99	89.38	90.10	0.3068

Mengacu pada Tabel 2, model BiGRU memperoleh ketepatan pelatihan sebanyak 90.99% serta ketepatan validasi ialah 89.38%. Perbedaan tingkat ketepatan antara data pelatihan dan data validasi berada pada rentang yang relatif rendah, sehingga kondisi ini mencerminkan tidak

terjadinya *overfitting* secara berarti pada model. Pada fase pengujian, capaian akurasi model tercatat sebesar 90,10% dengan nilai *loss* uji sebesar 0,3068. Temuan tersebut mengindikasikan bahwasanya model memperlihatkan performa yang konsisten serta mampu melakukan klasifikasi sentimen secara baik pada data yang belum terlibat pada proses pelatihan.

4.6. Evaluasi

Selanjutnya, kinerja kedua model dievaluasi untuk menilai efektivitasnya dalam mengklasifikasikan data uji. Pengukuran kinerja model dilakukan melalui pemanfaatan *confusion matrix* yang mencakup indikator akurasi, presisi, *recall*, dan f1-score. Gambaran hasil evaluasi tersebut merefleksikan tingkat kemampuan model dalam membedakan sentimen positif dan negatif pada keseluruhan data uji.

Tabel 3. Hasil Perbandingan Algoritma

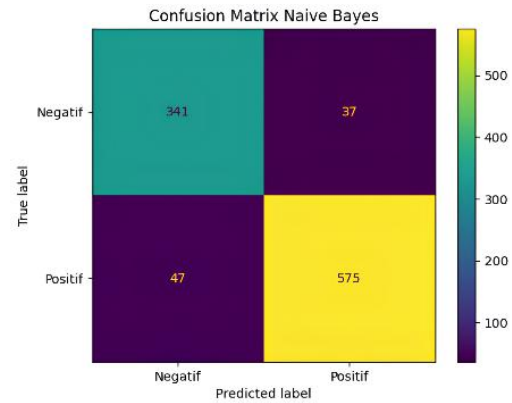
Model	Kelas	Precision	Recall	F1-Score
Naïve Bayes	Negatif	0.88	0.90	0.89
	Positif	0.94	0.92	0.93
BiGRU	Negatif	0.85	0.90	0.87
	Positif	0.94	0.90	0.92

Mengacu pada Tabel 3, perolehan atas pengkategorian memperlihatkan bahwasanya model *Naïve Bayes* memperoleh *recall* sebesar 92% bagi ulasan positif serta 90% bagi ulasan negatif, lewat *precision* tiap-tiap nya ialah 94% serta 88%, sekaligus F1-score 93% bagi ulasan positif serta 89% bagi ulasan negatif. Sedangkan, model BiGRU mencapai *recall* 90% bagi kedua kelas ulasan, dengan *precision* 94% bagi ulasan positif serta 85% bagi ulasan negatif, sekaligus F1-score 92% serta 87% bagi ulasan positif serta negatif. Kedua pemodelan memperlihatkan kinerja yang baik, namun *Naïve Bayes* lebih unggul dalam menjaga keseimbangan antara performa ulasan positif dan negatif, sedangkan BiGRU tetap kompetitif, khususnya pada *precision* ulasan positif.

Tabel 4. Hasil Akurasi

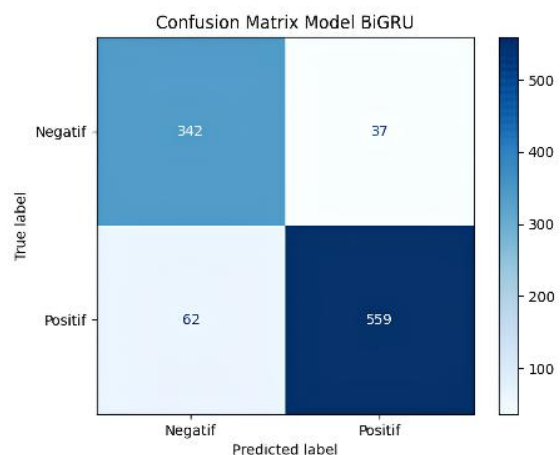
Model	Akurasi
Naïve Bayes	0.92
BiGRU	0.90

Berdasarkan Tabel 4, metode *Naïve Bayes* menghasilkan akurasi sebesar 92%, sedangkan BiGRU sebesar 90%. Keunggulan *Naïve Bayes* meliputi proses komputasi cepat, algoritma sederhana, serta kemampuan belajar efektif dari data terbatas, BiGRU tetap kompetitif melalui kemampuan menangkap pola urutan teks meskipun akurasinya sedikit lebih rendah



Gambar 4. Confusion Matrix Naïve Bayes

Confusion matrix dalam Gambar 4, memperlihatkan bahwasanya *Naïve Bayes* mengelompokkan 341 ulasan negatif serta 575 ulasan positif dengan benar, dengan 37 ulasan negatif dan 47 ulasan positif salah diklasifikasikan. Hasil ini memperlihatkan kemampuan model mengenali ulasan positif serta negatif dengan presisi.



Gambar 5. Confusion Matrix BiGRU

Dalam Gambar 5, *Confusion matrix* memperlihatkan bahwasanya BiGRU mengklasifikasikan 342 ulasan negatif serta 559 ulasan positif dengan benar, dengan 37 ulasan negatif dan 62 ulasan positif salah diklasifikasikan. Hasil ini memperlihatkan kemampuan model mengenali ulasan positif dan negatif dengan akurat.

5. KESIMPULAN DAN SARAN

Perolehan studi ini memperlihatkan bahwasanya algoritma *Naïve Bayes* serta *Bidirectional Gated Recurrent Unit* (BiGRU) mampu menangani klasifikasi sentimen ulasan pengguna aplikasi Gojek secara efektif. Proses penelitian mencakup pengumpulan data lewat *web scraping* di Google Play Store, diikuti langkah *preprocessing* yang meliputi *cleaning*, *case folding*, *tokenization*, *stopword removal*, dan *stemming*, serta pembobotan kata memanfaatkan TF-IDF bagi model *Naïve Bayes* dan tokenisasi, *padding*, serta *embedding* untuk BiGRU

sebelum dilakukan klasifikasi dan evaluasi menggunakan *confusion matrix*. Meskipun kedua algoritma memiliki akurasi yang relatif tinggi, yaitu 92% bagi *Naïve Bayes* serta 90% bagi BiGRU, ditemukan diferensiasi pada kapabilitas atas tiap-tiap pemodelan. *Naïve Bayes* memperlihatkan *recall* untuk ulasan positif sebesar 92% dan *precision* yang seimbang pada kedua kelas, sehingga klasifikasi yang dihasilkan lebih konsisten antara ulasan positif dan negatif. Sementara itu, BiGRU memiliki *precision* lebih tinggi untuk ulasan positif (94%), namun *recall* dan F1-score untuk ulasan negatif sedikit lebih rendah, memperlihatkan model ini lebih unggul dalam menangkap pola urutan teks tetapi kurang seimbang pada klasifikasi ulasan negatif. Perihal tersebut memperlihatkan bahwasanya *Naïve Bayes* lebih unggul pada efisiensi serta konsistensi pengkategorian pada data terbatas, sedangkan BiGRU tetap kompetitif dengan kemampuannya menangkap pola urutan teks meskipun akurasinya sedikit lebih rendah. Berdasarkan hasil penelitian, disarankan agar pengembang Gojek memanfaatkan analisis sentimen ulasan pengguna sebagai acuan untuk meningkatkan kualitas layanan, terutama dalam merespons keluhan dan pengalaman negatif dengan lebih cepat. Penelitian selanjutnya dapat mengeksplorasi algoritma klasifikasi lain, selayaknya *Logistic Regression*, *Random Forest*, atau pemodelan melalui *deep learning* seperti BERT, yang memiliki kemampuan menangani pola teks yang lebih kompleks.

DAFTAR PUSTAKA

- [1] D. A. Crystrie, S. Adhianur, and S. H. Sartika, "Pengaruh Inovasi Ojek menjadi Ojek Online pada Ekonomi Kreatif di Era Revolusi Industri 4.0 (Studi Kasus pada Masyarakat Kota Tasikmalaya)," vol. 2, no. 2, pp. 68–77, 2022.
- [2] R. A. Rahman, V. H. Pranatawijaya, and N. N. K. Sari, "Analisis Sentimen Berbasis Aspek pada Ulasan Aplikasi Gojek," vol. 4, no. 1, pp. 70–82, 2024.
- [3] Eskiyyaturoffikoh and R. R. Suryono, "Analisis sentimen aplikasi x pada google play store menggunakan algoritma naïve bayes dan support vector machine (svm)," vol. 9, no. 3, pp. 1408–1419, 2024.
- [4] N. C. B. Silalahia *et al.*, "ANALISIS SENTIMEN KEPUASAN PENGGUNA MEDIA SOSIAL DENGAN MENGGUNAKAN DATA MINING DAN MATLAB," vol. 7, no. 2, pp. 212–223, 2024.
- [5] S. Zhahira, O. A. Wiratama, S. Winarni, I. G. N. M. Jaya, and A. Chadidjah, "Analisis Sentimen Ulasan Pengguna Aplikasi Depok Single Window Menggunakan Metode Bidirectional Gate Recurrent Unit (BiGRU)," vol. 5, no. 1, pp. 44–53, 2024.
- [6] K. D. Indarwati and H. Februariyanti, "Analisis sentimen terhadap kualitas pelayanan aplikasi go-jek menggunakan metode naive bayes classifier," vol. 10, no. 1, 2023.
- [7] O. M. Nurfauzi, S. S. Hilabi, F. Nurapriani, and B. Huda, "Analisis Sentimen Grab Indonesia pada Ulasan Google Play Store Menggunakan Algoritma Naïve Bayes dan Support Vector Machine," vol. 11, no. 1, pp. 8–13, 2025.
- [8] V. N. Yudoyono, J. Maulana, A. R. Alfath, and M. A. Sihalo, "ANALISIS KOMPARATIF BILSTM DAN BIGRU DENGAN WORD EMBEDDING GLOVE TERHADAP SENTIMEN PUBLIK TENTANG COVID-19 DI TWITTER," vol. 13, no. 3, 2025.
- [9] N. Kadek, T. Ari, I. G. A. Gunadi, and I. M. G. Sunarya, "Analisis Sentimen Pelayanan Daring di Fakultas Teknik dan Kejuruan Universitas Pendidikan Ganesha Menggunakan Algoritma Naïve Bayes dan LSTM," vol. 4, no. 3, pp. 1120–1129, 2024.
- [10] P. W. Cahyo and U. S. Aesyi, "Perbandingan LSTM dengan Support Vector Machine dan Multinomial Naïve Bayes pada Klasifikasi Kategori Hoax," vol. 20, no. 2, pp. 23–29, 2023.
- [11] L. O. Lukmana, D. R. L, and P. Elisya, "APLIKASI THREADS INSTAGRAM DI PLAYSTORE MENGGUNAKAN ALGORITMA NAIVE BAYES," vol. 13, no. 2, 2025.
- [12] A. R. Putra and D. E. Ratnawati, "ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI MOBILE MENGGUNAKAN NAÏVE BAYES BERDASARKAN ULASAN PENGGUNA PLAYSTORE (STUDI KASUS : JCONNECT MOBILE)," vol. 12, no. 2, pp. 293–300, 2025, doi: 10.25126/jtiik.2025127556.
- [13] D. W. Bhatara and R. R. Suryono, "Analisis sentimen aplikasi bca mobile menggunakan algoritma naïve bayes dan support vector machine," vol. 9, no. 4, pp. 1907–1917, 2024.
- [14] M. N. Muttaqin and I. Kharisudin, "ANALISIS SENTIMEN APLIKASI GOJEK MENGGUNAKAN SUPPORT VECTOR MACHINE DAN K NEAREST NEIGHBOR," vol. 10, no. 2, 2021.
- [15] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, vol. 10, no. 1, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [16] Q. Zhu, X. Jiang, and R. Ye, "Sentiment Analysis of Review Text Based on BiGRU-Attention and Hybrid CNN," *IEEE Access*, vol. 9, pp. 149077–149088, 2021, doi: 10.1109/ACCESS.2021.3118537.
- [17] M. L. Akmal and D. Pernadi, "Analisis Sentimen Program Makan Siang Gratis Menggunakan Multinomial Naive Bayes," *J. Ilm. Sinus*, vol. 23, no. 1, pp. 117–126, 2025.