

## Analisis Review Pengguna Terhadap Fitur Baru Whatsapp Menggunakan LSTM

Rainanda Darya Saputra<sup>a</sup>, Tasrif Hasanuddin<sup>b</sup>, Ramdaniah<sup>c</sup>

Universitas Muslim Indonesia, Makassar, Indonesia

<sup>a</sup>rainandadarya@gmail.com, <sup>b</sup>tasrif.hasanuddin@umi.ac.id, <sup>c</sup>ramdaniah@umi.ac.id

Received: 16-08-2025 | Revised: 20-10-2025 | Accepted: 10-11-2025 | Published: 29-12-2025

### Abstrak

Sebagai salah satu platform komunikasi paling banyak digunakan, WhatsApp terus menghadirkan berbagai fitur baru untuk meningkatkan kenyamanan penggunanya. Namun, setiap pembaruan fitur seringkali disertai dengan beragam tanggapan dari pengguna, yang banyak dituangkan melalui ulasan di Google Play Store. Penelitian ini bertujuan untuk mengklasifikasikan sentimen dari ulasan-ulasan tersebut menggunakan metode Long Short-Term Memory (LSTM), sebuah arsitektur Recurrent Neural Network yang unggul dalam menangani data teks yang bersifat sekuensial. Data diperoleh melalui teknik web scraping dan diproses melalui tahapan pra-pemrosesan seperti tokenisasi, pembersihan teks, dan embedding kata. Model LSTM dirancang menggunakan kombinasi layer embedding, bidirectional LSTM, dan dense untuk menghasilkan klasifikasi sentimen ke dalam tiga kategori: positif, negatif, dan netral. Berdasarkan hasil pelatihan, model menunjukkan peningkatan akurasi yang signifikan serta penurunan loss yang stabil, mencapai akurasi validasi di atas 85%. Hasil confusion matrix menunjukkan bahwa model mampu membedakan ketiga jenis sentimen dengan akurasi tinggi dan kesalahan klasifikasi yang minim. Uji coba pada input kalimat baru juga menunjukkan kemampuan model dalam mengenali sentimen secara kontekstual dengan baik. Dengan demikian, pendekatan LSTM terbukti efektif dalam memahami dan menganalisis opini pengguna terhadap fitur-fitur baru WhatsApp.

Kata kunci: Analisis Teks, LSTM, Sentimen, Ulasan Aplikasi, WhatsApp.

### Pendahuluan

WhatsApp telah menjadi salah satu aplikasi komunikasi paling dominan di dunia, terus berinovasi dengan meluncurkan fitur-fitur baru untuk meningkatkan pengalaman pengguna. Setiap peluncuran fitur baru ini seringkali memicu beragam respons dan ulasan dari penggunanya [1]. Memahami sentimen pengguna ini sangat krusial bagi pengembang untuk mengevaluasi keberhasilan fitur, mengidentifikasi area perbaikan, dan merencanakan pengembangan di masa mendatang. Ulasan pengguna, yang seringkali disampaikan melalui platform seperti Google Play Store atau Apple App Store, merupakan sumber data berharga yang mencerminkan opini dan pengalaman langsung mereka. Oleh karena itu, diperlukan metode analisis yang canggih untuk mengolah data teks tidak terstruktur ini menjadi wawasan yang bermakna [2].

Penelitian ini akan memanfaatkan dua metode utama: analisis sentimen dan Long Short-Term Memory (LSTM). Analisis sentimen adalah teknik Natural Language Processing (NLP) yang bertujuan untuk mengekstrak, mengidentifikasi, dan mengukur polaritas emosi (positif, negatif, atau netral) dari teks [3], [4]. Sementara itu, LSTM adalah jenis arsitektur Recurrent Neural Network (RNN) yang dirancang khusus untuk memproses data sekuensial dan mengatasi masalah vanishing gradient yang sering terjadi pada RNN tradisional [5]. Kemampuan LSTM untuk mempelajari dependensi jangka panjang dalam data membuatnya sangat cocok untuk analisis teks, di mana konteks kata-kata sebelumnya sangat mempengaruhi makna kata-kata selanjutnya [6], [7].

Kombinasi analisis sentimen dengan LSTM menawarkan sinergi yang kuat dalam menganalisis ulasan pengguna WhatsApp. Analisis sentimen akan memberikan dasar untuk mengklasifikasikan polaritas ulasan, sedangkan LSTM akan meningkatkan akurasi klasifikasi tersebut dengan memahami nuansa kontekstual dalam teks [8]. Dengan memanfaatkan LSTM, kita dapat membangun model yang mampu mengenali pola-pola kompleks dalam ulasan pengguna, seperti penggunaan sarkasme atau ungkapan yang ambigu, yang seringkali sulit ditangkap oleh metode analisis sentimen berbasis leksikon sederhana [9]. Manfaat dari pendekatan ini adalah kemampuan untuk secara otomatis memproses volume besar ulasan pengguna secara efisien, memberikan pemahaman mendalam tentang penerimaan fitur baru, dan mengidentifikasi tren sentimen yang mungkin tidak terlihat melalui tinjauan manual [10].

Beberapa penelitian terdahulu telah menunjukkan efektivitas penggunaan LSTM dalam analisis sentimen pada berbagai domain, seperti ulasan produk, tweet, dan berita. Misalnya, penelitian Arif Fajar Pangestu, 2024 menunjukkan tingkat akurasi yang cukup tinggi, di atas 85% mengeksplorasi [11]. Demikian pula, penelitian oleh Dloifur Rohman Alghifari, 2022 berhasil menerapkan LSTM untuk klasifikasi sentimen ulasan terhadap layanan Grab Indonesia dengan akurasi sebesar 91% [12]. Penelitian sebelumnya menunjukkan bahwa penerapan LSTM merupakan salah satu metode yang cocok digunakan untuk melakukan analisis sentimen untuk penelitian ini. Kemampuannya untuk menangani dependensi jangka panjang dalam data sekuensial, seperti teks, memungkinkan model ini memahami konteks dan nuansa sentimen dengan lebih baik.

Berdasarkan latar belakang yang telah dijelaskan, tujuan utama dari penelitian ini adalah untuk mengembangkan model analisis sentimen berbasis LSTM yang mampu mengklasifikasikan ulasan pengguna terhadap fitur baru WhatsApp secara akurat. Lebih spesifik lagi, penelitian ini berupaya untuk mengevaluasi kinerja model LSTM dalam mengidentifikasi polaritas sentimen, baik itu positif, negatif, maupun netral, dari berbagai ulasan pengguna. Selain itu, melalui analisis sentimen ini, penelitian ini juga bertujuan untuk memberikan wawasan mendalam mengenai persepsi dan respons pengguna terhadap fitur-fitur baru yang diluncurkan oleh WhatsApp, sehingga dapat menjadi masukan berharga bagi pengembang dalam memahami kepuasan pengguna dan merencanakan pengembangan fitur selanjutnya.

## Metode

Penelitian ini mengimplementasikan *Long Short-Term Memory (LSTM)* sebagai arsitektur jaringan saraf tiruan utama untuk melakukan analisis sentimen terhadap ulasan pengguna terkait fitur baru *WhatsApp*. Pemilihan *LSTM* didasarkan pada kapabilitasnya yang unggul dalam memproses data sekuensial, seperti teks, serta kemampuannya untuk menangani dependensi jangka panjang dan mengatasi permasalahan *vanishing gradient* yang kerap ditemui pada *recurrent neural network (RNN)* konvensional [13], [14]. Model *LSTM* akan dilatih untuk mengklasifikasikan setiap ulasan pengguna ke dalam kategori sentimen yang telah ditentukan, yaitu positif, negatif, atau netral, dengan mengenali pola dan konteks semantik yang terkandung dalam teks [15].

Implementasi model ini akan dilakukan dengan memanfaatkan Python sebagai bahasa pemrograman utama. Python dipilih karena ketersediaan berbagai library dan *framework machine learning* yang robust, seperti *TensorFlow* atau *PyTorch*, yang memfasilitasi seluruh tahapan penelitian, mulai dari pra-pemrosesan data ulasan (termasuk *tokenization, embedding, dan noise reduction*), perancangan arsitektur jaringan *LSTM*, hingga proses pelatihan dan evaluasi model [16].

Untuk mengukur dan mengevaluasi kinerja model klasifikasi sentimen yang dibangun menggunakan *LSTM*. *Confusion Matrix* adalah sebuah tabel yang digunakan dalam machine learning untuk mengevaluasi kinerja dari suatu model klasifikasi, yang dimana memungkinkan kita untuk mengevaluasi akurasi, presisi, recall, dan F1-Score dari algoritma tersebut. *Confusion Matrix* akan digunakan sebagai alat evaluasi primer [17].

Akurasi merupakan metrik yang menghitung rasio antara jumlah prediksi yang benar (baik *True Positive* maupun *True Negative*) dengan total keseluruhan data [18]. Akurasi memberikan indikasi umum mengenai proporsi prediksi yang tepat oleh model. Akurasi dihitung sebagai rasio antara jumlah prediksi yang benar (baik positif maupun negatif) dengan keseluruhan dataset [19]. Perhitungan akurasi dapat dilakukan menggunakan persamaan 1.

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

Precision merupakan rasio antara *true positive (TP)* dengan total prediksi positif, yang mencakup *true positive (TP)* dan *false positive (FP)*. Perhitungan precision dapat dilakukan menggunakan persamaan 2.

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

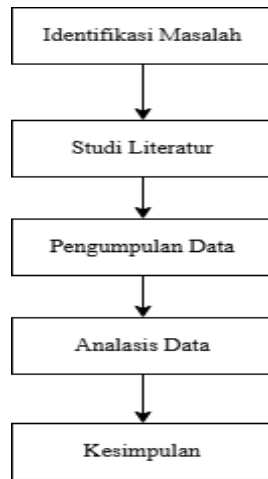
*Recall* adalah rasio antara jumlah prediksi benar positif dengan total data yang sebenarnya positif. Perhitungan *recall* dapat dilakukan menggunakan persamaan 3.

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

F1-Score adalah nilai rerata tertimbang dari precision dan recall. Rumus dalam menghitung F1-Score bias dilakukan dalam persamaan 4.

$$F1\ Score = 2 \times \frac{precision \times recall}{precision + recall} \tag{4}$$

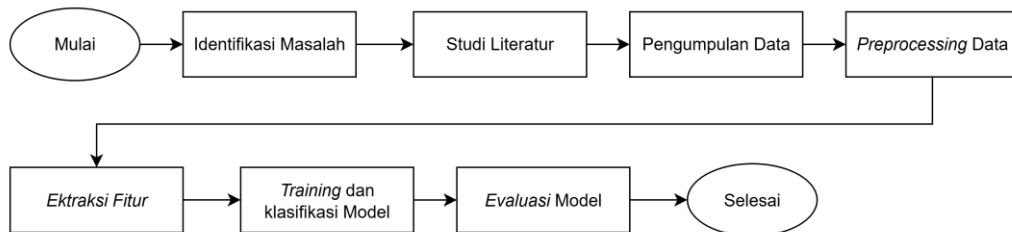
Berdasarkan persamaan tersebut, TP (*True Positive*) adalah prediksi yang benar sebagai kasus positif, TN (*True Negative*) adalah prediksi yang benar sebagai kasus negatif, FP (*False Positive*) adalah prediksi yang dianggap positif tetapi sebenarnya negatif, dan FN (*False Negative*) adalah prediksi yang dianggap negatif tetapi sebenarnya positif merupakan prediksi kasus positif ternyata negatif, dan FN (*False Negatif*) merupakan prediksi kasus negatif ternyata positif [20]. Tahapan penelitian yang dibangun ditujukan pada Gambar 1.



Gambar 1. Alur Penelitian

Tahapan penelitian dalam "Analisis Review Pengguna terhadap Fitur Baru *WhatsApp* Menggunakan *LSTM*" mencakup serangkaian langkah utama yang terstruktur. Proses penelitian diawali dengan Identifikasi Masalah, yakni menentukan fokus utama penelitian terkait ulasan pengguna terhadap fitur-fitur baru pada aplikasi *WhatsApp*. Selanjutnya, dilakukan Studi Literatur secara komprehensif guna memahami konsep analisis sentimen, arsitektur jaringan saraf *Long Short-Term Memory (LSTM)*, serta meninjau studi-studi sebelumnya yang relevan dalam bidang *NLP* dan analisis sentimen. Setelah itu, tahap Pengumpulan Data dilakukan dengan mengambil ulasan pengguna dari *Google Play Store* atau platform lain yang relevan sebagai dataset untuk dianalisis. Data ulasan yang terkumpul kemudian diproses dalam tahap Analisis Data, yang meliputi *preprocessing* teks (seperti *tokenization*, *stemming/lemmatization*, penghapusan *stop words*), representasi fitur teks (*word embedding*), serta penerapan dan pelatihan model *LSTM* untuk mengklasifikasikan sentimen ulasan ke dalam kategori positif, negatif, atau netral. Terakhir, pada tahap Kesimpulan, hasil analisis dievaluasi dan diinterpretasikan untuk memahami pola sentimen pengguna serta memberikan wawasan mengenai penerimaan fitur baru *WhatsApp* yang diteliti.

**Perancangan**



Gambar 2. Perancangan Penelitian

Rancangan penelitian analisis sentimen ini dimulai dengan Identifikasi Masalah untuk menentukan fokus fitur baru *WhatsApp* yang akan dianalisis, diikuti oleh Studi Literatur guna mendalami konsep analisis sentimen, arsitektur *LSTM*, dan metode pra-pemrosesan data teks. Selanjutnya, tahap Pengumpulan Data akan berfokus pada pengambilan ulasan pengguna dari platform aplikasi, yang kemudian akan menjalani Pra-pemrosesan Data (pembersihan, tokenisasi, *stop word removal*, *stemming*) untuk mempersiapkan teks. Tahap Ekstraksi Fitur akan melibatkan pembentukan *word embedding* sebagai representasi numerik data teks, yang menjadi

masuk ke dalam model Training & Klasifikasi Model LSTM yang dilatih untuk mengategorikan sentimen (positif, negatif, netral). Akhirnya, Evaluasi Model akan dilakukan menggunakan metrik seperti Akurasi, Presisi, Recall, F1-Score, dan Confusion Matrix untuk mengukur kinerja, yang hasilnya akan disimpulkan untuk memberikan wawasan tentang sentimen pengguna terhadap fitur baru WhatsApp

**Pemodelan**

**A. Scraping Data**

Dataset ini disusun dengan menerapkan teknik *data scraping* untuk menghimpun secara otomatis ulasan-ulasan yang diberikan oleh pengguna aplikasi WhatsApp dari berbagai platform digital publik. Proses pengumpulan data ini difokuskan secara spesifik untuk menyaring dan menganalisis umpan balik yang berkaitan langsung dengan penerimaan serta sentimen pengguna terhadap fungsionalitas atau fitur-fitur baru yang diperkenalkan. Hasilnya adalah sebuah kumpulan data komprehensif yang mencakup ragam ekspresi pengguna mulai dari opini tekstual, kritik, hingga saran yang secara otentik merefleksikan persepsi mereka terhadap inovasi pada aplikasi tersebut Berikut adalah 5 dataset teratas yang di gunakan:

```

5 Data Teratas:
                                review sentiment
0  Fitur polling di grup sangat membantu untuk me...  positif
1  Saya suka bagaimana sekarang bisa mengirim fil...  positif
2  Kualitas panggilan video semakin jernih setela...  positif
3  Kenapa fitur barunya membuat WhatsApp jadi ber...  negatif
4  Tampilan antarmukanya jadi aneh, lebih suka de...  negatif

Distribusi Sentimen:
sentiment
positif      182
negatif      162
netral       156
    
```

Gambar 3. Dataset

**B. Preprocessing Data**

*Preprocessing* Data merupakan Langkah awal dalam analisis teks yang bertujuan untuk mengubah data teks mentah menjadi format yang lebih terstruktur dan siap digunakan. Proses ini mencakup serangkaian tahapan dasar untuk membersihkan teks dari elemen-elemen yang tidak diperlukan atau kurang relevan, seperti tanda baca, angka, dan kata-kata umum yang tidak memiliki makna khusus dalam konteks analisis. Dengan demikian, data teks menjadi lebih bersih dan optimal untuk digunakan dalam tahap pemrosesan atau analisis *LSTM*. Berikut adalah hasil dari *Preprocessing* :

```

Contoh Data Setelah Preprocessing:
                                review
0  Fitur polling di grup sangat membantu untuk me...
1  Saya suka bagaimana sekarang bisa mengirim fil...
2  Kualitas panggilan video semakin jernih setela...
3  Kenapa fitur barunya membuat WhatsApp jadi ber...
4  Tampilan antarmukanya jadi aneh, lebih suka de...

                                cleaned_review
0  fitur polling grup membantu mengambil keputusan
1  suka mengirim file gb berguna pekerjaan
2  kualitas panggilan video jernih pembaruan
3  fitur barunya whatsapp berat not responding
4  tampilan antarmukanya aneh suka desain
    
```

Gambar 3. Hasil *Preprocessing* :

Gambar 3 menghasilkan teks ulasan yang lebih bersih dan terstruktur dengan menghapus komponen-komponen yang tidak diperlukan, seperti tanda baca, huruf besar, serta kata-kata umum yang tidak memberikan pengaruh signifikan terhadap analisis. Proses ini membuat teks menjadi lebih padat dan fokus pada inti pesan yang disampaikan oleh pengguna. Kalimat-kalimat yang sebelumnya panjang dan

mengandung kata-kata tambahan diubah menjadi lebih ringkas dan mudah dipahami. Tujuan dari penyederhanaan ini adalah untuk meningkatkan kualitas data sehingga lebih siap digunakan dalam tahapan analisis lanjutan, seperti klasifikasi atau analisis sentimen. Dengan data yang telah diproses ini, hasil analisis dapat menjadi lebih akurat dan bermakna.

C. Proses Pelatihan Model *LSTM*

Model yang digunakan dirancang menggunakan pendekatan Sequential dengan beberapa layer utama, antara lain layer Embedding, Bidirectional *LSTM*, dan *Dense*. Penambahan hidden layer dan penggunaan dropout bertujuan untuk meningkatkan kemampuan generalisasi model. Arsitektur ini diharapkan mampu mengklasifikasikan sentimen ke dalam tiga kategori dengan lebih baik. Berikut adalah hasil dari pelatihan model :

```
Epoch 1/12
13/13 - 18s - 1s/step - accuracy: 0.3725 - loss: 1.0963 - val_accuracy: 0.3600 - val_loss: 1.0864
Epoch 2/12
13/13 - 4s - 287ms/step - accuracy: 0.4175 - loss: 1.0811 - val_accuracy: 0.3600 - val_loss: 1.0685
Epoch 3/12
13/13 - 5s - 352ms/step - accuracy: 0.4000 - loss: 1.0601 - val_accuracy: 0.4400 - val_loss: 1.0367
Epoch 4/12
13/13 - 4s - 312ms/step - accuracy: 0.5550 - loss: 1.0104 - val_accuracy: 0.5200 - val_loss: 0.9662
Epoch 5/12
13/13 - 5s - 400ms/step - accuracy: 0.5575 - loss: 0.9494 - val_accuracy: 0.6800 - val_loss: 0.8694
Epoch 6/12
13/13 - 6s - 485ms/step - accuracy: 0.6675 - loss: 0.8265 - val_accuracy: 0.7900 - val_loss: 0.7380
Epoch 7/12
13/13 - 4s - 290ms/step - accuracy: 0.7425 - loss: 0.6958 - val_accuracy: 0.7400 - val_loss: 0.6467
Epoch 8/12
13/13 - 4s - 275ms/step - accuracy: 0.7825 - loss: 0.5897 - val_accuracy: 0.7900 - val_loss: 0.5175
Epoch 9/12
13/13 - 7s - 573ms/step - accuracy: 0.8700 - loss: 0.4372 - val_accuracy: 0.8600 - val_loss: 0.4348
Epoch 10/12
13/13 - 9s - 716ms/step - accuracy: 0.9000 - loss: 0.3484 - val_accuracy: 0.8300 - val_loss: 0.4112
Epoch 11/12
13/13 - 4s - 291ms/step - accuracy: 0.9075 - loss: 0.2778 - val_accuracy: 0.8700 - val_loss: 0.3215
Epoch 12/12
13/13 - 5s - 392ms/step - accuracy: 0.9350 - loss: 0.2237 - val_accuracy: 0.9000 - val_loss: 0.2932
```

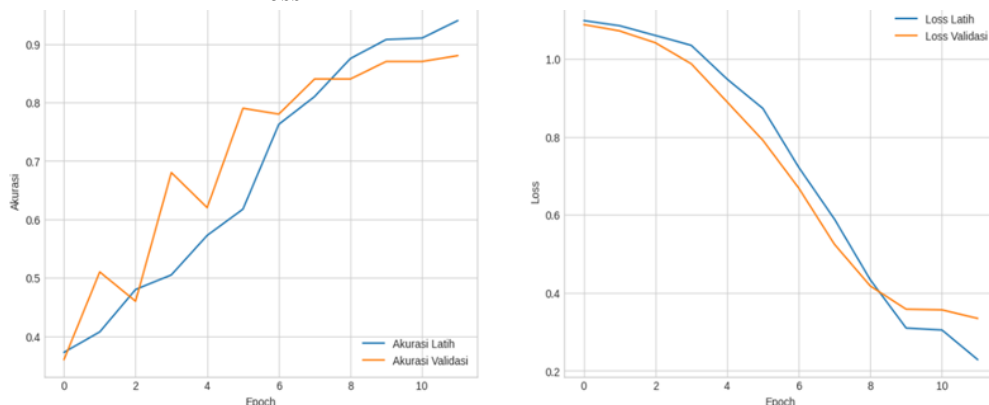
Gambar 4. Output Model dan Hasil Pelatihan Selama 12 Epoch

Gambar 4 model dilatih selama 12 epoch menggunakan *optimizer Adam* dan *loss function categorical\_crossentropy*. Berdasarkan hasil pelatihan, model menunjukkan peningkatan akurasi secara bertahap dari epoch ke epoch. Akurasi validasi meningkat dari 0.36 menjadi 0.90 pada akhir epoch, sedangkan nilai *loss* menurun secara signifikan dari 1.08 menjadi 0.29. Hal ini mengindikasikan bahwa model berhasil mempelajari data dengan baik tanpa mengalami *overfitting*.

D. Evaluasi Kinerja Model

Evaluasi terhadap kinerja model dilakukan untuk mengetahui seberapa efektif model dalam melakukan klasifikasi data. Pada penelitian ini, proses evaluasi mencakup metrik akurasi, nilai *loss*, dan visualisasi confusion matrix. Selain itu, performa model selama proses pelatihan dan validasi juga dianalisis untuk mengetahui kestabilan dan kemampuan generalisasi model.

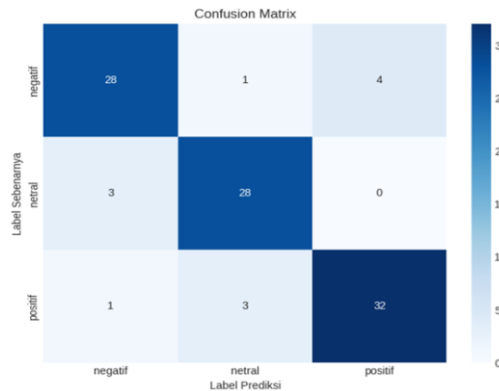
1. Visualisasi Akurasi dan *Loss*



Gambar 5. Grafik Akurasi & Grafik *Loss*

Gambar 5 menampilkan grafik perkembangan nilai akurasi dan loss model selama pelatihan dalam beberapa *epoch*. Visualisasi ini berguna untuk menilai sejauh mana model mengalami peningkatan performa seiring waktu. Grafik Akurasi pada data pelatihan menunjukkan tren peningkatan yang konsisten, dimulai dari sekitar 0,37 dan mencapai lebih dari 0,95. Sementara itu, akurasi pada data validasi juga meningkat, meskipun mengalami sedikit fluktuasi. Hal ini menunjukkan bahwa model mampu mempelajari pola data dengan baik tanpa menunjukkan gejala *overfitting* yang signifikan. Grafik Loss: Grafik loss memperlihatkan penurunan kesalahan prediksi (*loss*) baik pada data pelatihan maupun validasi. Penurunan yang stabil mengindikasikan bahwa model belajar secara efektif dan mampu mengurangi *error* seiring bertambahnya *epoch*.

2. *Confusion Matrix*



Gambar 6. *Confusion Matrix*

Evaluasi klasifikasi dilakukan dengan menggunakan *confusion matrix* yang ditampilkan pada Gambar 6. Model dikembangkan untuk mengklasifikasikan data ke dalam tiga kelas, yaitu *negatif*, *netral*, dan *positif*. Hasil klasifikasi menunjukkan Untuk kelas negatif, sebanyak 28 data berhasil diprediksi dengan benar, sedangkan sisanya diprediksi keliru sebagai *netral* (1 data) dan *positif* (4 data). Pada kelas netral, terdapat 28 data yang diklasifikasikan dengan benar, dan 3 data lainnya salah diklasifikasikan sebagai *negatif*. Untuk kelas positif, model menunjukkan performa terbaik dengan 32 prediksi benar, dan hanya 4 data yang salah diklasifikasikan ke kelas lain.

Secara umum, *confusion matrix* menunjukkan bahwa model memiliki performa yang cukup baik dan mampu membedakan antar kelas dengan tingkat akurasi yang tinggi. Distribusi kesalahan prediksi juga tidak menunjukkan dominasi kesalahan pada salah satu kelas.

3. Analisis Hasil Evaluasi

Berdasarkan hasil evaluasi yang dilakukan, model terbukti memiliki kemampuan yang cukup andal dalam melakukan klasifikasi data. Nilai akurasi validasi yang tinggi, yaitu di atas 85%, dan nilai loss yang terus menurun menunjukkan bahwa model memiliki generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya. Selain itu, tingkat kesalahan klasifikasi yang rendah pada confusion matrix menandakan bahwa model mampu meminimalkan kesalahan prediksi dengan baik dalam skenario multi-kelas.

E. Uji Coba Model Dengan Kalimat Baru

Setelah model selesai dilatih dan dievaluasi, langkah selanjutnya adalah melakukan pengujian terhadap kalimat-kalimat baru yang tidak termasuk dalam data latih maupun validasi. Tujuan dari pengujian ini adalah untuk mengamati kemampuan model dalam mengklasifikasikan sentimen secara nyata terhadap masukan teks bebas. Pengujian dilakukan dengan menggunakan tiga contoh kalimat baru dengan konteks yang berbeda, yaitu positif, negatif, dan netral. Berikut adalah hasil pengujian:

```

--- Uji Coba Model ---
1/1 ----- 0s 314ms/step
Kalimat Input: 'Fitur untuk mengedit chat sangat membantu sekali, terima kasih whatsapp!'
Sentimen Prediksi: POSITIF
Probabilitas (negatif, netral, positif): [0.01147197 0.00588685 0.98264116]
-----
1/1 ----- 0s 58ms/step
Kalimat Input: 'Update baru ini jelek sekali, membuat aplikasi jadi sangat lambat.'
Sentimen Prediksi: NEGATIF
Probabilitas (negatif, netral, positif): [0.98796296 0.00720113 0.00483589]
-----
1/1 ----- 0s 55ms/step
Kalimat Input: 'Saya tidak merasakan ada perubahan apa-apa.'
Sentimen Prediksi: NETRAL
Probabilitas (negatif, netral, positif): [0.00847699 0.99019146 0.00133145]

```

Gambar 6. Output Model Dengan Kalimat Baru

Gambar 6 Model yang telah dibangun kemudian diuji menggunakan beberapa kalimat baru yang tidak termasuk dalam data pelatihan maupun validasi, guna melihat sejauh mana kemampuan model dalam mengklasifikasikan sentimen secara nyata. Hasil uji coba menunjukkan bahwa model mampu mengidentifikasi sentimen dengan sangat baik. Kalimat bernada positif seperti "*Fitur untuk mengedit chat sangat membantu sekali, terima kasih whatsapp!*" berhasil diklasifikasikan sebagai positif dengan tingkat probabilitas mencapai 98,26%. Begitu pula kalimat dengan nada keluhan seperti "*Update baru ini jelek sekali, membuat aplikasi jadi sangat lambat.*" dikenali sebagai negatif dengan akurasi prediksi sebesar 98,79%. Sementara itu, kalimat yang bersifat netral seperti "*Saya tidak merasakan ada perubahan apa-apa.*" diprediksi sebagai netral dengan keyakinan hingga 99,01%. Kemampuan model dalam membedakan ketiga jenis sentimen ini menunjukkan bahwa model memiliki performa yang kuat dan dapat digunakan untuk mengklasifikasikan opini atau tanggapan pengguna dalam situasi nyata dengan tingkat kepercayaan yang tinggi.

## Kesimpulan

Hasil penelitian ini memperlihatkan bahwa algoritma *Long Short-Term Memory (LSTM)* mampu secara efektif digunakan dalam mengklasifikasikan sentimen pengguna terhadap fitur terbaru pada aplikasi *WhatsApp*. Model yang dibangun berhasil menunjukkan performa yang memuaskan, dengan tingkat akurasi validasi lebih dari 85% dan penurunan nilai loss yang stabil selama proses pelatihan. Evaluasi melalui *confusion matrix* memperlihatkan bahwa model mampu mengidentifikasi dengan tepat ketiga label sentimen: positif, negatif, dan netral, dengan kesalahan klasifikasi yang relatif rendah. Selain itu, model juga menunjukkan kinerja yang baik saat diuji menggunakan kalimat-kalimat baru di luar dataset pelatihan, yang membuktikan bahwa model memiliki tingkat generalisasi yang tinggi. Dengan hasil tersebut, dapat disimpulkan bahwa metode *LSTM* sangat tepat digunakan untuk memahami persepsi dan respon pengguna terhadap pembaruan fitur dalam aplikasi digital seperti *WhatsApp*.

## Daftar Pustaka

- [1] A. M. Nur, "Penggunaan Aplikasi Whatsapp Dalam Media Komunikasi di Perguruan Tinggi Pada Masa Pandemi Covid-19 (Studi Kasus: Universitas Al-Wasliyah Labuhan Batu)," *Communication & Social Media*, vol. 1, no. 2, pp. 42–48, Nov. 2021, doi: 10.57251/csm.v1i2.318.
- [2] Tania Puspa Rahayu Sanjaya, Ahmad Fauzi, and Anis Fitri Nur Masruriyah, "Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine," *INFOTECH: Jurnal Informatika & Teknologi*, vol. 4, no. 1, pp. 16–26, Jun. 2023, doi: 10.37373/infotech.v4i1.422.
- [3] M. Rosihan Anwar and N. Mufidah, "Istikhdam al-Dhakhā' al-Iṣṭinā'ī (AI) fi Ta'lim Ṭullāb Ṣaff al-Takhaṣṣuṣ fi Baḥṡ al-Lughah al-'Arabiyyah," *Al-Kafaah: Journal of Arabic Language and Linguistics Education (ALLE)*, pp. 1–17, Dec. 2024, doi: 10.52491/alle.vi.186.
- [4] A. Amelia, L. N. Hayati, and H. Darwis, "Analisis Sentimen Masyarakat Terhadap Sistem Pembayaran MyPertamina dengan Metode Random Forest, SVM, dan Naïve Bayes," *LINIER: Literatur Informatika dan Komputer*, vol. 1, no. 1, pp. 28–44, Nov. 2024, doi: 10.33096/linier.v1i1.2269.
- [5] L. Kristiana and D. Miyanto, "Penambahan Parameter PM2.5 dalam Prediksi Kualitas Udara : Long Short Term Memory," *Multimedia Artificial Intelligent Networking Database (MIND)*, vol. 8, no. 2, pp. 188–202, 2023.

- [6] E. H. A. Prastyo, "Deteksi berita hoax dengan pendekatan Lexicon Based dan LSTM," Universitas Islam Negeri Maulana Malik Ibrahim, 2024.
- [7] M. Fajar Abdillah and K. Kusnawi, "Comparative Analysis of Long Short-Term Memory Architecture for Text Classification," *ILKOM Jurnal Ilmiah*, vol. 15, no. 3, pp. 455–464, Dec. 2023, doi: 10.33096/ilkom.v15i3.1906.455-464.
- [8] K. S. Witanto, N. A. Sanjaya ER, A. E. Karyawati, I. G. A. G. A. Kadyanan, I. K. G. Suhartana, and L. G. Astuti, "Implementasi LSTM Pada Analisis Sentimen Review Film Menggunakan Adam Dan RMSprop Optimizer," *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, vol. 10, no. 4, p. 351, Jun. 2022, doi: 10.24843/JLK.2022.v10.i04.p05.
- [9] A. Saputra, R. C. Sigitta Hariyono, and N. M. Saraswati, "Analisis Sentimen Pengguna Aplikasi MyPertamina Menggunakan Algoritma Bidirectional Long Short Term Memory," *Jurnal Eksplora Informatika*, vol. 13, no. 2, pp. 156–163, Mar. 2024, doi: 10.30864/eksplora.v13i2.973.
- [10] Putri Angraini Aziz, S. B. Nur Ilahi, Sumiarni Moka, and A. M. Sajiah, "Penerapan Hadoop untuk Analisis Sentimen Berbasis Big Data pada Ulasan Aplikasi Transportasi Online," *SATESI: Jurnal Sains Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 51–60, Apr. 2025, doi: 10.54259/satesi.v5i1.4051.
- [11] A. F. Pangestu, B. Rahmat, and A. N. Sihananto, "Analisis Sentimen Pada Media Sosial X Terhadap Implementasi Kurikulum Merdeka Menggunakan Metode FastText Dan Long Short-Term Memory (LSTM)," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 4, pp. 2271–2280, Nov. 2024, doi: 10.29100/jupi.v9i4.5665.
- [12] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 12, no. 2, pp. 89–99, Sep. 2022, doi: 10.34010/jamika.v12i2.7764.
- [13] M. Aasya Aldin Islamy and P. Pandu Adikara, "Analisis Sentimen IMDB Movie Reviews menggunakan Metode Long Short-Term Memory dan FastText," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 9, pp. 4106–4115, 2022.
- [14] F. I. Putri, A. P. Wibawa, and L. H. Collante, "Refining the Performance of Indonesian-Japanese Bilingual Neural Machine Translation Using Adam Optimizer," *ILKOM Jurnal Ilmiah*, vol. 16, no. 3, pp. 271–282, Dec. 2024, doi: 10.33096/ilkom.v16i3.2467.271-282.
- [15] I. Kurniasari, AchmadAchmad Arif Alfin2, and E. Widodo, "Implementasi Long Short-Term Memory (LSTM) dan Word Embedding Model pada Analisis Sentimen Layanan Uang Elektronik Ovo dan Link Aja," *INFORMASI (Jurnal Informatika dan Sistem Informasi)*, vol. 15, no. 2, pp. 237–246, Nov. 2023, doi: 10.37424/informasi.v15i2.273.
- [16] Z. Annisa and B. S. S. Ulama, "Analisis Sentimen Data Ulasan Pengguna Aplikasi 'PeduliLindungi' pada Google Play Store Menggunakan Metode Naïve Bayes Classifier Model Multinomial," *Jurnal Sains dan Seni ITS*, vol. 11, no. 6, 2023, doi: 10.12962/j23373520.v11i6.94064.
- [17] N. Made *et al.*, "Analisis Sentimen Berbahasa Inggris Dengan Metode Lstm Studi Kasus Berita Online Pariwisata Bali English Sentiment Analysis Using The Lstm Method Case Study Of Bali Tourism Online News," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 11, pp. 1325–1334, 2024, doi: 10.25126/jtiik.2024118792.
- [18] K. Kariyamin, Muh. I. Alyakin, and L. O. Alyandi, "Menganalisis Ulasan Mobile Legends: Analisis Kinerja Berdasarkan Opini Pengguna dengan Naive Bayes," *Buletin Sistem Informasi dan Teknologi Islam*, vol. 6, no. 1, pp. 1–8, Feb. 2025, doi: 10.33096/busiti.v6i1.2475.
- [19] S. Wulandari and F. N. Hasan, "Analisis Sentimen Masyarakat Indonesia Terhadap Pengalaman Belanja Thrifting Pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, pp. 768–776, 2024.
- [20] Y. Ansori and K. F. H. Holle, "Perbandingan Metode Machine Learning dalam Analisis Sentimen Twitter," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 10, no. 4, p. 429, Dec. 2022, doi: 10.26418/justin.v10i4.51784.