

Metode Deep Learning LSTM dalam Analisis Sentimen Aplikasi PeduliLindungi

Siti Mutmainah^{1*}, Khairunnas², Khairunnisa³

¹Universitas Muhammadiyah Bima, Indonesia

* Email Korespondensi: siti.mutmainah.id19@gmail.com

Abstrak: Penyebaran pandemi virus corona pada awal tahun 2020 menciptakan ketegangan global, mendorong upaya pencegahan dan pengendalian di seluruh dunia, termasuk Indonesia. Respons pemerintah terhadap situasi ini melibatkan berbagai langkah proaktif, salah satunya adalah pengembangan aplikasi PeduliLindungi. Aplikasi ini berfungsi sebagai platform yang memungkinkan masyarakat berkontribusi dengan berbagi data lokasi saat beraktivitas. Tujuannya adalah mempermudah pelacakan riwayat kontak dengan individu yang terinfeksi Covid-19 di Indonesia. Kendati memiliki dampak positif dalam pengendalian penyebaran virus, semakin meluasnya penggunaan aplikasi PeduliLindungi memunculkan beragam pandangan di kalangan pengguna, yang tercermin dalam umpan balik pada *Google Play Store*. Untuk mengelola beragam data umpan balik berupa teks tersebut, pendekatan analisis sentimen muncul sebagai solusi. Dalam konteks ini, penelitian ini mengusulkan penerapan metode *Long Short-Term Memory* (LSTM) untuk menganalisis sentimen umpan balik aplikasi PeduliLindungi. Dengan mencapai akurasi sebesar 92,51%, dengan 69,3% aspek negatif, 18,3% positif, dan 12,4% netral, hasil analisis ini memberikan wawasan yang berharga bagi pengembang aplikasi dan pemangku kepentingan lainnya. Diharapkan hasil penelitian ini dapat digunakan sebagai panduan untuk meningkatkan kualitas dan efektivitas aplikasi PeduliLindungi guna memberikan dampak positif yang lebih besar dalam mengatasi tantangan kesehatan masyarakat.

Kata Kunci: LSTM, PeduliLindungi, Analisis Sentimen, Deep Learning, Covid-19

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



1. Pendahuluan

Pada tahun 2020 awal, penyebaran wabah virus corona terjadi di hampir seluruh negara di dunia, termasuk Indonesia. Kecepatan penyebaran dan penularan virus corona menimbulkan kepanikan dan ketakutan yang sangat besar di kalangan penduduk dunia [1]. Beberapa upaya telah dilakukan oleh Pemerintah, diantaranya penetapan kebijakan protokol kesehatan, menutup akses ke tempat umum, dan pembatasan sosial. Meski demikian, hal-hal tersebut dianggap kurang efektif, sehingga pemerintah memutuskan untuk menerapkan metode pelacakan kontak. Pada awalnya, pelacakan kontak dilakukan secara manual, di mana individu yang terinfeksi virus memberikan informasi mengenai orang-orang serta lokasi mereka melakukan kontak fisik. Namun, penerapan pelacakan manual ini dianggap sebagai pemborosan waktu dan sumber daya [2], [3].

Di era digital saat ini, peran teknologi informasi sangat signifikan dalam membentuk berbagai aspek kehidupan masyarakat. Sebagai contoh konkret dari dampak tersebut, terlihat pada adopsi aplikasi berbasis *mobile* sebagai alat yang mendukung upaya pencegahan dan pengendalian penyebaran penyakit, terutama

di tengah pandemi global virus corona (Covid-19). Pemerintah Indonesia melakukan upaya pengendalian penyebaran Covid-19 dengan berbagai pendekatan, salah satunya melibatkan penggunaan aplikasi Android yang dapat diunduh melalui *Google Play Store*. Aplikasi tersebut adalah PeduliLindungi, berfungsi sebagai *platform* untuk penanganan dan partisipasi masyarakat dalam berbagi data lokasi ketika berpergian. Tujuan utama aplikasi ini adalah memfasilitasi pelacakan riwayat kontak dengan individu yang terinfeksi Covid-19 di Indonesia. Dampaknya, masyarakat menjadi lebih aktif dalam menyampaikan pendapat melalui aplikasi PeduliLindungi yang tersedia di *Google Play Store* [4].

Aplikasi PeduliLindungi telah diunduh oleh lebih dari lima puluh juta pengguna, data ini diperoleh dari total unduhan di *platform Google Play Store*. Berdasarkan informasi dari Administrasi Kependudukan (Adminduk) pada Juni 2021, jumlah penduduk Indonesia mencapai 272.229.372 jiwa. Oleh karena itu, setidaknya 5% dari total penduduk Indonesia telah menggunakan aplikasi ini selama masa pandemi Covid-19 [5]. Aplikasi PeduliLindungi membuktikan kebermanfaatannya bagi masyarakat dalam menjalankan berbagai aktivitas, karena mampu menyediakan informasi seperti penelusuran (*tracing*), pelacakan (*tracking*), dan pembatasan ruang gerak (*fencing*). Meski demikian, semakin meluasnya penggunaan aplikasi PeduliLindungi, muncul beragam pandangan di kalangan masyarakat yang menggunakannya. Sudut pandang masyarakat terkait penggunaan aplikasi PeduliLindungi perlu dipahami dengan baik [6]. Ragam pandangan atau seterusnya disebut sebagai umpan balik oleh pengguna PeduliLindungi pada *Google Play Store* dapat menjadi acuan bagi pemerintah untuk mengetahui apa yang dipikirkan dan diinginkan oleh masyarakat, serta mengetahui alasan mengapa aplikasi PeduliLindungi kurang diminati atau mengapa aplikasi tersebut banyak digunakan. Informasi tersebut dapat digunakan sebagai referensi untuk masa depan guna meningkatkan kualitas dari aplikasi PeduliLindungi [7]. Salah satu metode untuk mengkategorikan umpan balik adalah dengan membaginya menjadi positif, netral, atau negatif. Dengan mengelompokkan umpan balik tersebut, dapat diperoleh informasi mengenai fakta dan opini masyarakat terhadap produk, merek, layanan, politik, atau topik tertentu, dalam hal ini aplikasi PeduliLindungi. Bidang ini dikenal sebagai analisis sentimen dan termasuk dalam ranah *Natural Language Processing* (NLP) [8].

Analisis sentimen merupakan proses klasifikasi apakah suatu blok teks bersifat positif, negatif, atau netral. Ini merupakan eksplorasi kontekstual kata-kata yang mencerminkan sentimen sosial terhadap suatu merek dan turut membantu perusahaan menilai apakah produk yang dihasilkan akan diminati di pasar atau tidak. Tujuan analisis sentimen adalah mengevaluasi pendapat orang dengan cara yang mendukung perkembangan bisnis. Fokusnya tidak hanya pada polaritas (positif, negatif, dan netral) tetapi juga pada aspek emosi (senang, sedih, marah, dan sebagainya). Analisis sentimen adalah teknik yang memungkinkan untuk mengidentifikasi dan menentukan apakah data mencerminkan emosi positif, negatif, atau netral [9].

Penelitian terkait analisis sentimen untuk aplikasi PeduliLindungi sebelumnya telah dilakukan, diantaranya menggunakan beberapa metode seperti, *random forest* [5], *Support Vector Machine* (SVM) [3], *Naïve Bayes* dan *K-Nearest Neighbor* [6], serta *K-means-SVM* dan *DBSCAN-SVM* [10].

Studi mengenai analisis sentimen terhadap umpan balik Aplikasi PeduliLindungi khususnya pada *App Store*, yang menggunakan metode *random forest* dengan kedalaman pohon (*tree*) sebesar 65 dan jumlah pohon (*tree*) sebanyak 400, berhasil mencapai tingkat akurasi sebesar 72% [5]. Sedangkan analisis sentimen yang melibatkan penggunaan metode *Support Vector Machine* (SVM) dengan penerapan seleksi fitur *chi-square*, hasilnya mencapai tingkat akurasi sebesar 93% [3]. Penelitian lain terkait dengan evaluasi sentimen terhadap Aplikasi PeduliLindungi telah dilaksanakan dengan melibatkan 321 dataset. Hasil penelitian ini menunjukkan tingkat akurasi sebesar 70,46% untuk metode *Naïve Bayes* dan 73,33% untuk metode *K-Nearest Neighbor* [6]. Tingkat akurasi yang sangat tinggi, yakni sekitar 99,6% dan 99,8%, berhasil dicapai melalui penerapan metode *K-means-SVM* dan *DBSCAN-SVM* dalam penelitian ini. Sumber data yang digunakan berasal dari *platform Google Playstore* dan *App Store* [10]. Disamping metode-metode yang telah dijelaskan sebelumnya, terdapat satu metode analisis sentimen lain yang sering digunakan, yaitu *Long Short-Term Memory* (LSTM).

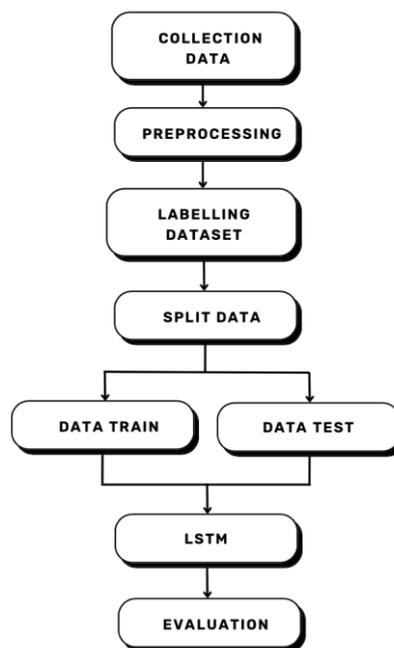
LSTM adalah suatu bentuk arsitektur *Recurrent Neural Network* (RNN) yang diformulasikan untuk menyimpan nilai-nilai yang telah diperoleh pada periode sebelumnya. Arsitektur LSTM terdiri dari tiga gerbang pengatur aliran, yaitu *input gate*, *forget gate*, dan *output gate*. *Input gate* berfungsi mengatur masukan data baru ke dalam memori, *forget gate* mengontrol berapa lama nilai-nilai tertentu tetap disimpan dalam memori, dan *output gate* mengontrol sejauh mana nilai yang tersimpan [11]. Sebuah penelitian telah dilaksanakan untuk melakukan analisis sentimen terhadap umpan balik Hotel Indonesia dengan

mengintegrasikan metode LSTM dan Word2Vec. Hasilnya menunjukkan tingkat akurasi sebesar 85.96% [11]. Penelitian lain mencatat tingkat akurasi yang signifikan, mencapai 85%, ketika menggunakan LSTM untuk melakukan analisis sentimen berbasis teks [12]. Selain itu, suatu penelitian lainnya mengadakan analisis sentimen terhadap umpan balik film dengan mengimplementasikan beberapa model, termasuk *Logistic Regression*, SVM dengan kernel linear, MLP, DNN, dan LSTM. Performa terbaik tercatat pada metode LSTM dengan tingkat akurasi sebesar 88,46% [13].

Dengan merujuk pada permasalahan yang telah diidentifikasi dan kajian literatur yang telah dilakukan, penelitian ini mengajukan pendekatan dengan mengadopsi metode *Long Short-Term Memory* (LSTM) dalam melakukan analisis sentimen terhadap umpan balik yang diterima terkait aplikasi PeduliLindungi di platform *Google Play Store*.

2. Metodologi Penelitian

Penelitian ini melibatkan serangkaian langkah pemodelan yang akan diterapkan. Gambar 1 menyajikan diagram air yang secara visual menampilkan tahapan pelaksanaan penelitian.



Gambar. 1 Tahapan Penelitian

2.1 Pengumpulan Data

Penelitian ini berfokus pada umpan balik pengguna terhadap aplikasi PeduliLindungi yang terdapat pada *platform Google Play Store*. Proses pengumpulan data menggunakan teknik *web scraping*, dan data yang terkumpul disimpan dalam format CSV untuk dikelola dan dianalisis lebih lanjut. Data umpan balik yang terkumpul mencakup rentang waktu Juni hingga Desember 2021, dengan total 27.263 entri. Rentang waktu tersebut dipilih untuk merepresentasikan persepsi pengguna secara komprehensif selama periode yang signifikan ini. Proses pengumpulan data yang sistematis diharapkan dapat memberikan gambaran yang mendalam mengenai dinamika sentimen terhadap aplikasi PeduliLindungi.

2.2 Preprocessing

Preprocessing merupakan tahap penting yang bertujuan untuk meningkatkan kualitas data yang digunakan dalam penelitian ini. Langkah-langkah *preprocessing* melibatkan *text cleaning*, *case folding*, *tokenisasi*, *stopword removal*, dan *stemming*. Tujuan utama dari langkah-langkah ini adalah untuk menghilangkan *noise* dan ketidaksempurnaan dari *dataset*, serta menstandarisasi kata-kata untuk

memastikan bahwa data dapat diproses secara efisien oleh komputer. Tujuan dari langkah-langkah *preprocessing* ini adalah untuk membangun kerangka dasar yang kuat untuk algoritme pemodelan yang akan digunakan pada tahap selanjutnya dari penelitian ini. Hal ini akan memungkinkan analisis data yang akurat dan efektif.

2.3 Labelling

Pada penelitian ini menganalisis data sentimen dengan menggunakan *leksikon* sebagai metode analisis umpan balik aplikasi PeduliLindungi. Penggunaan *leksikon* membantu penentuan sentimen secara otomatis di setiap umpan balik, sehingga memperkuat dasar untuk eksplorasi lebih lanjut mengenai tanggapan pengguna terhadap aplikasi kesehatan. *Lexicon* model berisi kamus yang memiliki polaritas untuk setiap kata untuk mengekstrak titik data dari setiap teks, output yang dihasilkan adalah skor yang digunakan untuk menentukan label sentimen.

2.4 Pemodelan LSTM

Model yang digunakan yaitu LSTM jenis dari metode RNN dan modifikasi RNN, LSTM memanfaatkan memori jangka panjang dan jangka pendek sehingga memiliki efek *skip connection* yang dapat menangani *vanishing gradien*. LSTM memiliki 4 layers: *Forget gate*, *Input gate*, *Cell state* dan *Output gate*.

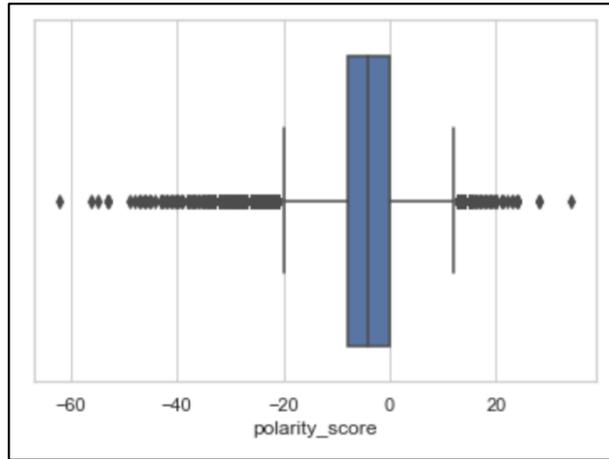
2.5 Evaluasi

Evaluasi kinerja model menggunakan *confusion matrix*. Matriks ini menampilkan empat komponen utama: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Metrik evaluasi, seperti akurasi, presisi, recall, dan F1-score, dapat dihitung dari matriks ini. Interpretasi hasil evaluasi model akan didasarkan pada nilai yang dihasilkan oleh *confusion matrix*. Akurasi model akan menunjukkan kemampuannya untuk memprediksi sentimen dengan benar. Selain itu, presisi, recall, dan F1-score akan memberikan wawasan lebih lanjut tentang kinerja model untuk setiap kategori sentimen.

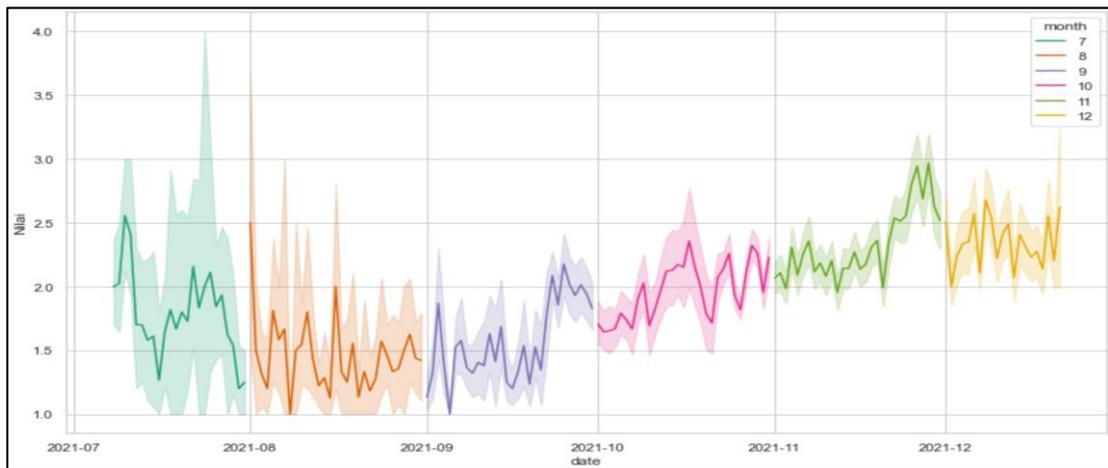
3. Hasil dan Pembahasan

3.1 Exploratory Data Analysis

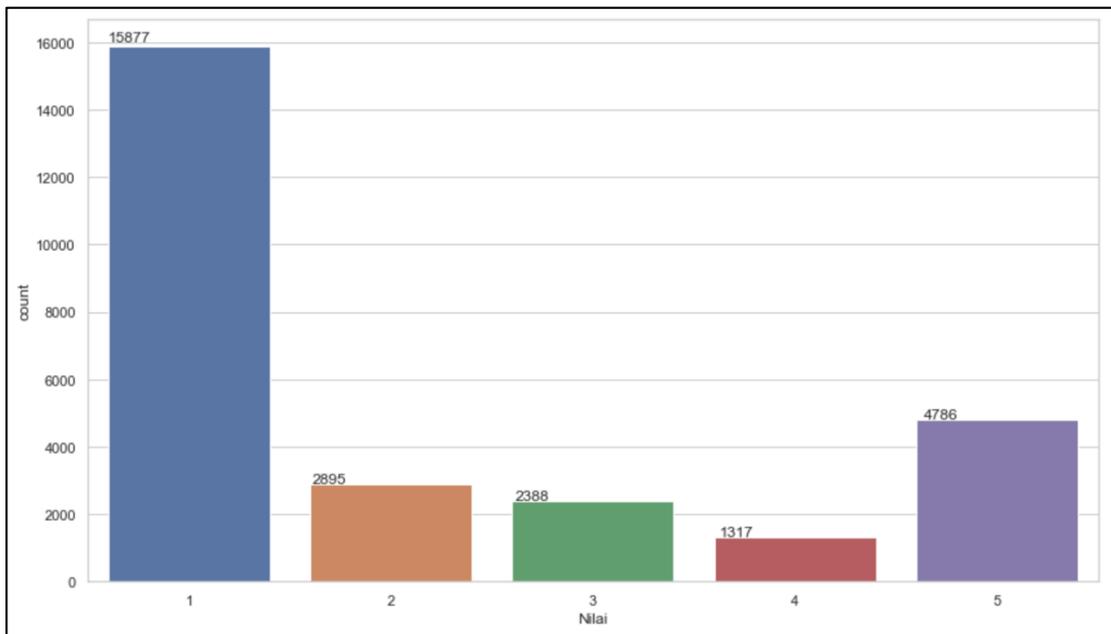
Exploratory Data Analysis (EDA) untuk mengeksplorasi dan menjelaskan hasil analisis umpan balik aplikasi PeduliLindungi. EDA merupakan langkah penting dalam memahami karakteristik utama dari dataset. EDA menyajikan visualisasi awal data, termasuk histogram dan *boxplot*. *Boxplot* akan memberikan pandangan yang lebih rinci tentang distribusi data, dengan fokus pada variabel yang relevan dengan penggunaan aplikasi. Grafik frekuensi akan digunakan untuk menggambarkan pola penggunaan aplikasi dari waktu ke waktu, yang menunjukkan seberapa sering aplikasi PeduliLindungi digunakan setiap bulannya. Analisis ini bertujuan untuk mengidentifikasi tren penggunaan yang mungkin terkait dengan acara khusus atau kampanye tertentu. Analisis rating aplikasi akan mencakup distribusi nilai rating yang diberikan oleh pengguna. Melalui visualisasi, penelitian ini akan mengidentifikasi tren umum rating pengguna terhadap aplikasi PeduliLindungi dan menganalisis hubungan antara penggunaan aplikasi dan rating yang diberikan. Pada Gambar 2 hingga Gambar 4 menyajikan *boxplot*, grafik frekuensi penggunaan bulanan, dan rating aplikasi PeduliLindungi.



Gambar. 2 Boxplot



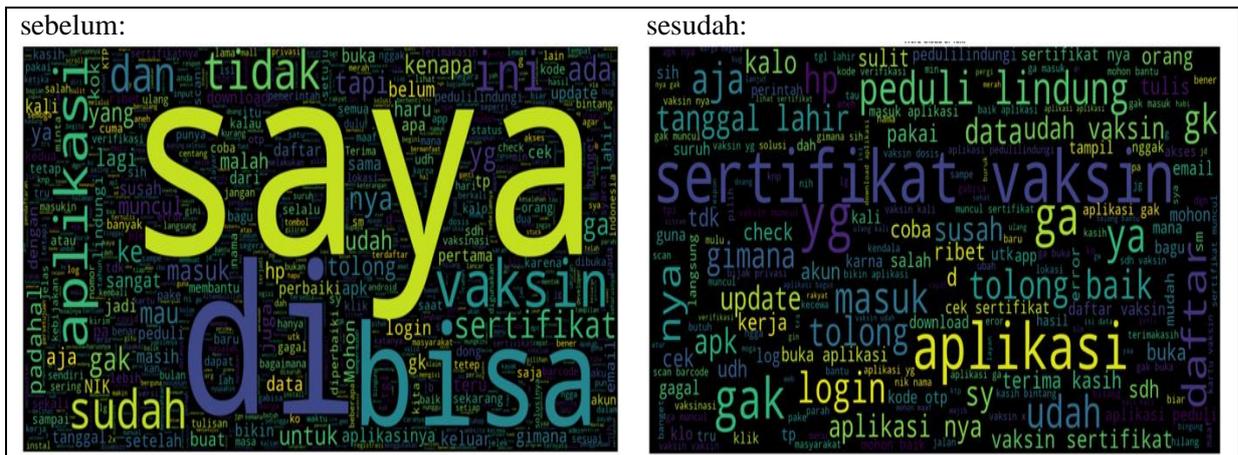
Gambar. 3 Frekuensi Pengguna PeduliLindungi



Gambar. 4 Rattng Pengguna PeduliLindungi

3.2 Preprocessing

Preprocessing merupakan tahap penting dalam persiapan data sebelum analisis sentimen. Untuk memberikan gambaran yang lebih jelas, hasil *wordcloud* sebelum dan sesudah *preprocessing* di sajikan pada Gambar 5. Sebelum *preprocessing*, *wordcloud* umpan balik pengguna aplikasi PeduliLindungi menampilkan sejumlah besar kata-kata yang berpotensi tidak relevan atau derau. Kata-kata ini dapat mengganggu analisis sentimen dan menghasilkan representasi yang kurang akurat dari sentimen yang sebenarnya. Namun, setelah *preprocessing*, *wordcloud* umpan balik pengguna menggambarkan kata-kata yang lebih fokus dan relevan yang lebih baik dalam menyampaikan sentimen yang diungkapkan. Kemampuan model untuk mengenali pola sentimen telah ditingkatkan dengan menghapus kata-kata yang tidak relevan dan *preprocessing* teks.



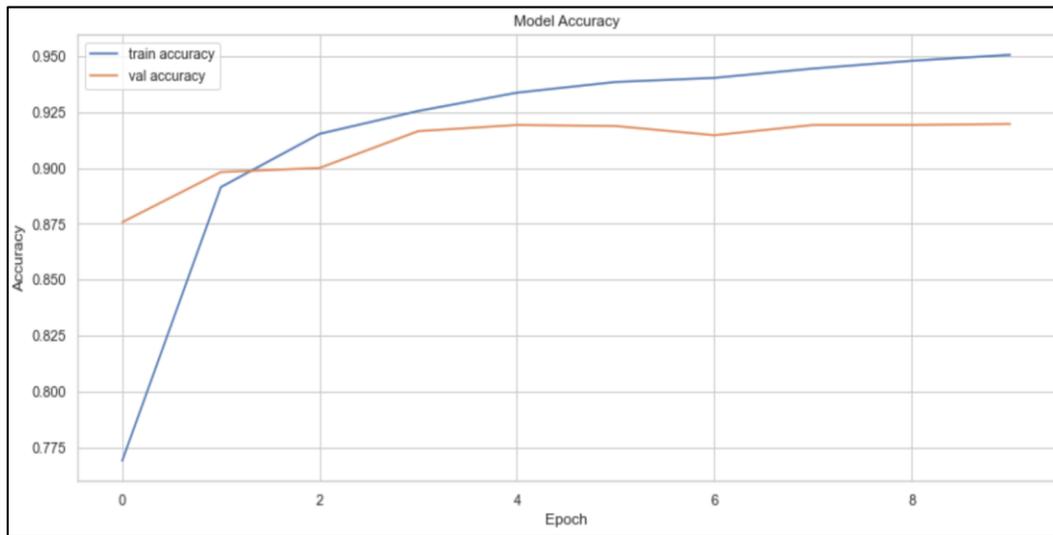
Gambar. 5 Preprocessing Data

3.3 Hasil Labelling Sentimen

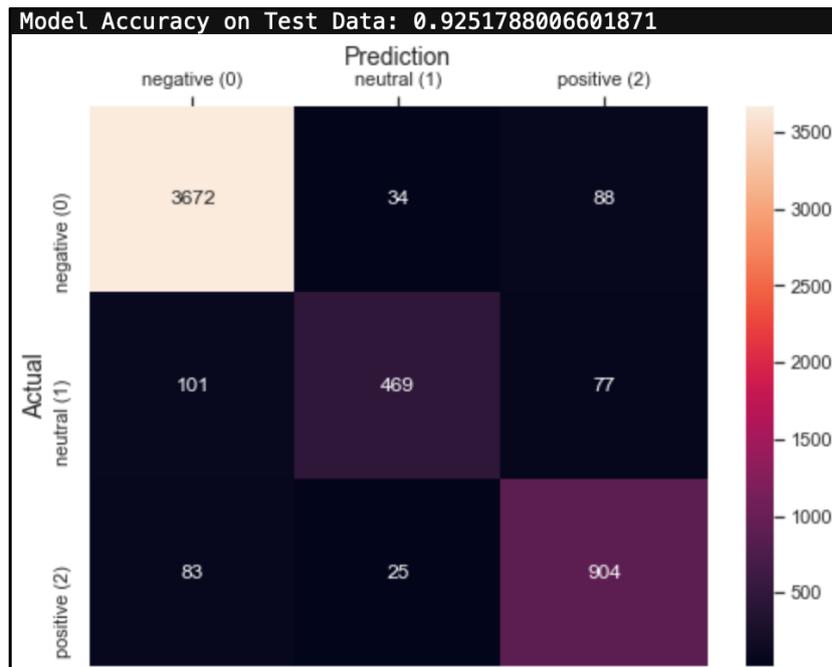
Hasil analisis sentimen menggunakan leksikon menunjukkan distribusi label yang signifikan pada dataset umpan balik PeduliLindungi. Dataset ini berisi 18.890 sentimen negatif, 5.001 sentimen positif, dan 3.372 sentimen netral. Analisis ini memberikan pemahaman yang lebih dalam tentang tanggapan pengguna terhadap aplikasi. Gambar 2 menunjukkan distribusi label sentimen pengguna aplikasi PeduliLindungi.

3.4 Hasil Pemodelan LSTM

Gambar 3 menyajikan distribusi sentimen pengguna dan pemodelan sentimen untuk aplikasi PeduliLindungi menggunakan *Long Short-Term Memory* (LSTM). Proses pelatihan menggunakan 21.810 sampel, sementara pengujian dilakukan pada 5.453 sampel untuk mengevaluasi kinerja model. Hasilnya menunjukkan tingkat akurasi sebesar 92,51%. Sebelum melakukan pemodelan, dataset dipartisi menjadi dua subset: data pelatihan dan data pengujian. Proses pelatihan dilakukan pada data pelatihan yang terdiri dari 21.810 sampel untuk memungkinkan model memahami pola dan konteks sentimen yang diekspresikan dalam umpan balik pengguna. Setelah menyelesaikan pelatihan, model dievaluasi pada *dataset* yang berbeda yang terdiri dari 5.453 sampel untuk menilai kemampuannya secara keseluruhan dalam memprediksi sentimen.



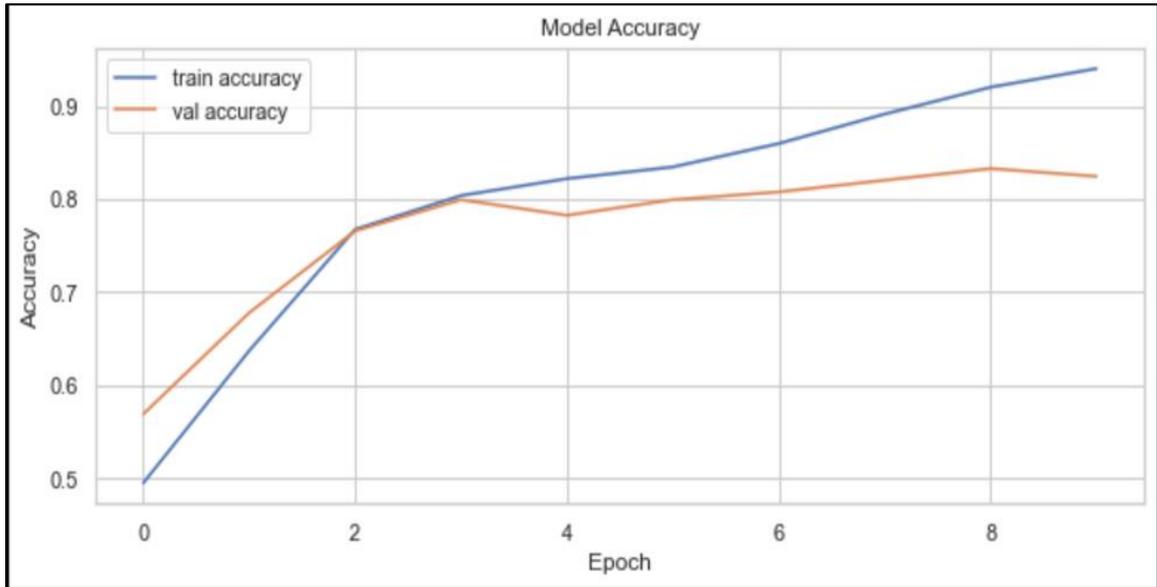
Gambar. 8 Sentimen Pengguna PeduliLindungi



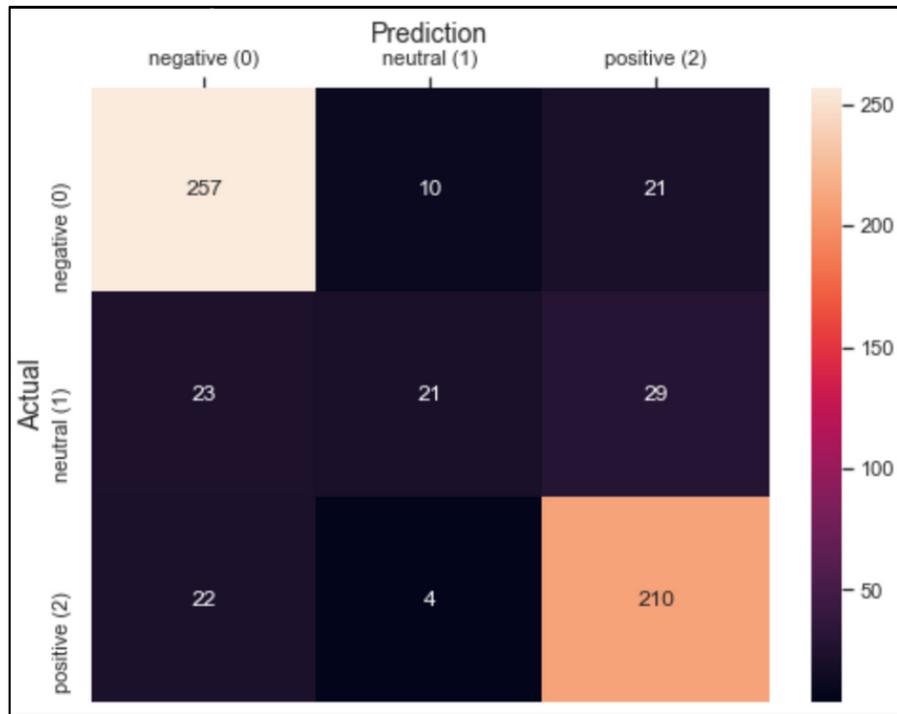
Gambar. 9 Sentimen Pengguna PeduliLindungi

Confusion matrix merepresentasikan hasil pengujian model dengan membandingkan kelas aktual dengan kelas yang diprediksi oleh model LSTM. Matriks ini menggambarkan kemampuan model untuk

mengklasifikasikan sentimen ke dalam kategori positif, negatif, dan netral. Model ini mencapai akurasi keseluruhan sebesar 92,51%, yang mengindikasikan kinerja yang baik dalam memprediksi sentimen dari umpan balik pengguna.



Gambar. 10 Sentimen Pengguna PeduliLindungi



Gambar. 11 Sentimen Pengguna PeduliLindungi

4. Kesimpulan

Hasil penelitian menunjukkan bahwa model LSTM dapat secara akurat mengidentifikasi dan memahami konteks sentimen yang kompleks dalam untuk aplikasi PeduliLindungi. Akurasi LSTM adalah 92,51%, akurasi yang tinggi menunjukkan kemampuan model dalam menangani variasi ekspresi bahasa dan nuansa sentimen yang ada di dalam dataset. Model LSTM dapat menjadi model yang efektif untuk menganalisis sentimen dengan tingkat akurasi yang tinggi, sehingga membantu pengembang untuk merespon umpan balik dari pengguna dengan lebih efektif. Kesimpulan ini menjadi dasar untuk rekomendasi dan tindakan lebih lanjut untuk meningkatkan pengalaman pengguna aplikasi PeduliLindungi. Penelitian ini memberikan informasi tentang bagaimana respon pengguna terhadap aplikasi PeduliLindungi. Hasil ini dapat digunakan untuk membuat keputusan strategis dalam pengembangan aplikasi dan meningkatkan layanan kepada pengguna. Perbaikan dapat diarahkan pada area-area kritis untuk mencapai pengalaman pengguna yang lebih positif.

5. REFERENSI

- [1] I. W. Sudiarsa and I. G. B. Wiraditya, "Analisis Usability Pada Aplikasi Peduli Lindungi Sebagai Aplikasi Informasi Dan Tracking Covid-19 Dengan Heuristic Evaluation," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 3, no. 2, pp. 354–364, Dec. 2020, doi: 10.31539/intecom.v3i2.1901.
- [2] K. Ahmad *et al.*, "Sentiment Analysis of Users' Reviews on COVID-19 Contact Tracing Apps with a Benchmark Dataset," Mar. 2021, [Online]. Available: <http://arxiv.org/abs/2103.01196>
- [3] C. Chairunnisa, I. Ernawati, and M. M. Santoni, "Klasifikasi Sentimen Ulasan Pengguna Aplikasi PeduliLindungi di Google Play Menggunakan Algoritma Support Vector Machine dengan Seleksi Fitur Chi-Square," *Informatik : Jurnal Ilmu Komputer*, vol. 18, no. 1, p. 69, Aug. 2022, doi: 10.52958/iftk.v17i4.4594.
- [4] A. Noviriandini, H. Hermanto, and Y. Yudhistira, "KLASIFIKASI SUPPORT VECTOR MACHINE BERBASIS PARTICLE SWARM OPTIMIZATION UNTUK ANALISA SENTIMEN PENGGUNA APLIKASI PEDULILINDUNGI," *JIKA (Jurnal Informatika)*, vol. 6, no. 1, p. 50, Apr. 2022, doi: 10.31000/jika.v6i1.5681.
- [5] M. Reza, U. Pulungan, D. E. Ratnawati, and B. Rahayudi, "Analisis Sentimen Ulasan Aplikasi PeduliLindungi dengan Metode Random Forest," 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [6] A. P. Wibowo, W. Darmawan, and N. Amalia, "KOMPARASI METODE NAÏVE BAYES DAN K-NEAREST NEIGHBOR TERHADAP ANALISIS SENTIMEN PENGGUNA APLIKASI PEDULILINDUNGI," *IC-Tech*, vol. 17, no. 1, pp. 18–23, Apr. 2022, doi: 10.47775/icttech.v17i1.234.
- [7] V. Christanti Mawardi and E. Darmaja, "Logistic Regression Method for Sentiment Analysis Application on Google Playstore," *International Journal of Application on Sciences, Technology and Engineering*, vol. 1, no. 1, pp. 241–247, Feb. 2023, doi: 10.24912/ijaste.v1.i1.241-247.
- [8] D. Abror, "Analisis Sentimen Review Aplikasi PeduliLindungi Menggunakan Seleksi Fitur Information Gain Berbasis SVM," *Indonesian Journal on Software Engineering (IJSE)*, vol. 9, no. 1, pp. 1–8, 2023, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ijse>
- [9] V. Dudhankar, N. Sen, A. Langde, and V. Kupade, "GOOGLE PLAYSTORE REVIEW SENTIMENT ANALYSIS," Jun. 2022. [Online]. Available: www.irjmets.com
- [10] A. T. 'Ain and S. Supatmi, "Comparison analysis of K-means-SVM and DBSCAN-SVM algorithm in determining sentiment users of PeduliLindungi," 2023, p. 030002. doi: 10.1063/5.0176473.
- [11] P. F. Muhammad, R. Kusumaningrum, and A. Wibowo, "Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews," *Procedia Comput Sci*, vol. 179, pp. 728–735, 2021, doi: 10.1016/j.procs.2021.01.061.

- [12] G. S. N Murthy, S. Rao Allu, B. Andhavarapu, M. Bagadi, and M. Belusonti, “Text based Sentiment Analysis using LSTM; Text based Sentiment Analysis using LSTM,” May 2020. [Online]. Available: www.ijert.org
- [13] J. D. Bodapati, N. Veeranjanyulu, and S. Shaik, “Sentiment analysis from movie reviews using LSTMs,” *Ingenierie des Systemes d’Information*, vol. 24, no. 1, pp. 125–129, 2019, doi: 10.18280/isi.240119.