SOSCIED Vol. 8 No. 2 November 2025
p-ISSN : 2622-8866
e-ISSN : 2721-9550

IMPLEMENTASI ALGORITMA *DEEP LEARNING* UNTUK ANALISIS SENTIMEN PENGGUNA PLATFORM PENDIDIKAN

IMPLEMENTATION OF DEEP LEARNING ALGORITHMS FOR SENTIMENT ANALYSIS OF EDUCATIONAL PLATFORM USERS

Anjis Sapto Nugroho[1], Eko Prasetyo[2], Adhi Priyanto[3], Daniel Alfa Puryono[4]

[1]STMIK AKI Pati
Jln. Kamandowo No.13 Pati
Jawa Tengah, Indonesia
anjie.nugros@gmail.com

[2]STMIK AKI Pati
Jln. Kamandowo No.13 Pati
Jawa Tengah, Indonesia
1pras1406@gmail.com

[3]STMIK AKI Pati
Jln. Kamandowo No.13 Pati
Jawa Tengah, Indonesia
adhi.stmikaki@gmail.com

[4]STMIK AKI Pati
Jln. Kamandowo No.13 Pati
Jawa Tengah, Indonesia
danielsempurna@gmail.com

ABSTRACT

More than 3.5 million Indonesian educators had used the Merdeka Mengajar Platform by 2024. Comparing this figure to the 3.37 million from the previous academic year, there has been an increase of almost 3.85%. However, an investigation is required to determine the reasons why the application's use has not yet achieved the anticipated goal number of users. This study does sentiment analysis on evaluations of the Merdeka Mengajar platform using Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM). The benefits of RNN and LSTM in processing sequential data, especially in text processing for sentiment analysis, led to their selection. The purpose of this study is to solve the difficulties in determining if users' sentiments on the platform are favorable or negative. Important steps in the research technique include preprocessing, data cleaning, and employing FastText embedding to convert text into numerical vectors. Then, using patterns in the text data, RNN and LSTM models are used to forecast sentiment. The study's findings demonstrate that the LSTM model can, with an estimated accuracy of 93.58%, identify long-term associations in sequential data. The RNN model, on the other hand, produces a lesser accuracy of 91.70%. Particularly in text data with intricate temporal circumstances, the LSTM model performs better at accurately classifying sentiment. By better understanding customer opinions and input on the Merdeka Mengajar platform, this study helps platform developers improve the quality of their services.

Keywords : *Deep Learning, Sentimen Analisis, LSTM, FastText, Akurasi*

1. PENDAHULUAN

Kemajuan teknologi informasi dan komunikasi telah berdampak signifikan pada berbagai sektor, termasuk pendidikan[1]. Di Indonesia, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) meluncurkan platform Merdeka Mengajar (PMM) yang dirancang untuk meningkatkan kualitas pengajaran yang diberikan oleh guru di Indonesia sebagai bagian dari upaya mendukung pelaksanaan program Merdeka[2]. Sejak diluncurkan pada tahun 2022 hingga 2024, platform ini telah digunakan oleh lebih dari 3,5 juta guru di Indonesia, meningkat 3,85% dari tahun sebelumnya[3]. Namun, jumlah tersebut masih jauh dari target yang diharapkan oleh semua guru di Indonesia. Penggunaan aplikasi ini belum optimal karena berbagai tantangan, seperti kurangnya akses ke daerah terpencil dan kesulitan teknis yang dihadapi oleh guru[4]. Ini menunjukkan kesenjangan antara tujuan platform dan kebutuhan pengguna. Pada akhirnya, situasi ini menimbulkan pertanyaan tentang seberapa baik platform ini memenuhi kebutuhan pengguna, yang dapat diungkapkan melalui analisis sentimen ulasan pengguna.

Mengingat hal ini, analisis sentimen ulasan pengguna menjadi alat penting untuk menilai pengalaman pengguna guna memahami pendapat dan komentar pengguna mengenai layanan teknologi, seperti platform Merdeka Mengajar[5], serta memberikan saran berbasis data untuk peningkatan. Analisis ini dapat mengkategorikan komentar pengguna sebagai positif atau negatif[6], menggunakan teknik pemrosesan bahasa alami (NLP)[7]. Ini membantu pengembang meningkatkan kualitas layanan dan memberikan wawasan tentang faktor-faktor yang memengaruhi adopsi platform oleh guru. Pemrosesan data sekuensial yang kompleks merupakan tantangan bagi teknik pembelajaran mesin

tradisional seperti Naive Bayes dan Support Vector Machine[8]. Kendala-kendala ini diatasi oleh kemampuan jaringan saraf berulang (RNN)[9]. dan memori jangka pendek panjang (LSTM) sebagai algoritma pembelajaran mendalam[10] untuk merekam pola temporal dan asosiasi antar kata. Dalam berbagai tugas analisis sentimen, LSTM terbukti lebih efektif daripada RNN karena mekanisme memorinya yang canggih, terutama saat berhadapan dengan data sekuensial yang kompleks dan data dalam jumlah besar[11].

Karena sebagian besar penelitian berfokus pada platform internasional atau aplikasi komersial, masih sedikit literatur tentang penggunaan algoritma RNN dan LSTM dalam analisis sentimen aplikasi pendidikan di Indonesia, meskipun penggunaannya luas [12]. Untuk menutup kesenjangan ini, penelitian ini menggunakan algoritma RNN dan LSTM berbasis FastText embeddings, yang mampu menangkap asosiasi semantik kata, terutama dalam bahasa Indonesia [13]. Melalui penggunaan embedding FastText, yang jarang diteliti dalam konteks pendidikan, penelitian ini bertujuan untuk meningkatkan presisi prediksi sentimen dalam ulasan, terutama bagi pengguna Platform Merdeka Mengajar.

Masalah penelitian yang dibahas dalam penelitian ini adalah bagaimana meningkatkan akurasi analisis sentimen [14]. pada ulasan pengguna PMM untuk mencapai hasil yang lebih baik dan bagaimana hasil analisis ini dapat digunakan untuk memahami persepsi pengguna dan meningkatkan kualitas layanan platform [15]. Untuk menjawab pertanyaan ini, penelitian mengambil pendekatan komprehensif yang mencakup pembersihan data, praproses, dan penggunaan input FastText untuk merepresentasikan teks sebagai vektor digital [16]. Model RNN dan LSTM kemudian diterapkan untuk memprediksi sentimen berdasarkan pola dalam data penambangan teks [17]. Target audiens utama untuk penelitian ini mencakup akademisi yang tertarik mengembangkan metode NLP dan pembelajaran mendalam, serta praktisi di bidang teknologi pendidikan, khususnya pengembang aplikasi pendidikan. Bagi akademisi, penelitian ini berkontribusi pada pemahaman tentang kelebihan dan keterbatasan algoritma RNN dan LSTM serta mencakup pengembangan ilmu NLP dan pembelajaran mendalam. Bagi pengembang platform, hasil penelitian ini dapat digunakan sebagai panduan untuk meningkatkan kualitas layanan dan pengalaman pengguna dalam aplikasi pendidikan di Indonesia.

2. KAJIAN PUSTAKA

Studi sebelumnya tentang analisis sentimen cuitan COVID-19 menggunakan metode RNN dilakukan oleh [18], dan berfokus pada ulasan tentang virus di Twitter. Selain menjelaskan penggunaan lapisan embedding untuk merepresentasikan kata sebagai vektor numerik yang dapat diproses oleh model pembelajaran mendalam, penelitian ini juga menjelaskan prosedur praproses seperti penghapusan tanda baca, normalisasi huruf, tokenisasi, dan stemming. Kinerja klasifikasi RNN dinilai dalam penelitian ini, dengan tingkat akurasi 90%. Temuan investigasi menunjukkan bahwa metode ini dapat secara akurat mendeteksi tren sentimen dan menawarkan pemahaman mendalam tentang bagaimana masyarakat bereaksi terhadap pandemi melalui saluran digital. Kinerja klasifikasi RNN dinilai dalam penelitian ini, dengan tingkat akurasi 90%. Temuan investigasi menunjukkan bahwa metode ini dapat secara akurat mendeteksi tren sentimen dan menawarkan pemahaman mendalam tentang bagaimana masyarakat bereaksi terhadap pandemi melalui saluran digital.

Efektivitas algoritma Long Short-Term Memory (LSTM) dan Naive Bayes dalam analisis sentimen diuji dalam penelitian oleh [19], khususnya dalam kaitannya dengan respons publik terhadap kebijakan "New Normal" selama pandemi COVID-19. Dengan 1.590 postingan dari Twitter, penelitian ini menggunakan sampel yang seimbang yang berfokus pada sentimen positif dan negatif. Dengan akurasi, presisi, dan recall sebesar 83,33% dibandingkan dengan 82% untuk Naive Bayes, hasilnya menunjukkan bahwa LSTM berkinerja lebih baik daripada Naive Bayes dan juga lebih cepat. Signifikansi analisis sentimen untuk memahami opini publik dan mengarahkan keputusan kebijakan ditekankan oleh penelitian ini.

Untuk mengatasi masalah data yang tidak seimbang dalam klasifikasi sentimen, penelitian berikut menggunakan metode Jaringan Saraf Berulang (RNN) untuk melakukan analisis sentimen terhadap evaluasi pengguna aplikasi Shopee Indonesia [20]. Teknik Oversampling Minoritas Sintetik (SMOTE) dan Tautan Tomek digunakan bersamaan untuk persiapan data dalam penelitian ini, yang meningkatkan ukuran kinerja seperti akurasi (80%) dan skor F1 (88,1%). Hasil ini menyoroti betapa pentingnya mengontrol data yang tidak seimbang untuk meningkatkan kualitas hasil kategorisasi sentimen.

Menurut sebuah studi yang menyelidiki penerapan teknik pembelajaran mendalam, Word2Vec dan Long Short-Term Memory (LSTM) digunakan dalam upaya menganalisis sentimen ulasan film. Ini membahas kesulitan dalam menguraikan dokumen teks panjang dan menunjukkan seberapa baik LSTM menangani jenis urutan ini. Menggunakan kumpulan data 25.000 ulasan film, penelitian ini menguji beberapa dimensi vektor kata menggunakan pendekatan Skip-Gram. Akurasi terbaik sebesar 88,17% diperoleh dengan ukuran vektor kata 100, sedangkan metode CBOW mencapai akurasi 87,68% pada dimensi 200. Signifikansi dimensi vektor kata dan kemungkinan penyesuaian parameter tambahan untuk meningkatkan akurasi disorot dalam penelitian ini [21].

Studi lain oleh Subowo et al., (2022) [22] meningkatkan analisis sentimen komentar menggunakan Bi-LSTM dan pendekatan Corpus Text untuk pelabelan otomatis pada aplikasi belanja online yang menerima pembayaran cicilan. Model Bi-LSTM untuk klasifikasi sentimen dan embedding FastText untuk representasi kata digabungkan dalam penelitian ini untuk menganalisis data platform. Hasil penelitian menunjukkan bahwa pendekatan model Bi-LSTM dengan teks korpus yang diberi label otomatis dapat mencapai akurasi 81% jika dibandingkan dengan metode tradisional. Ini memberikan data yang lebih akurat tentang bagaimana perasaan orang terhadap aplikasi belanja cicilan online.

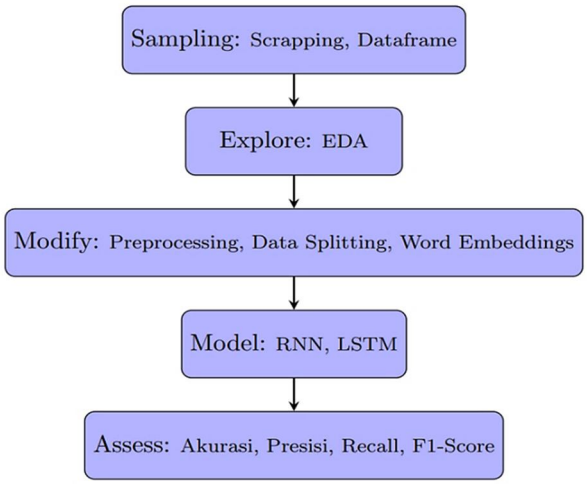
Penelitian yang disebutkan di atas menunjukkan bahwa analisis sentimen dengan algoritma RNN dan LSTM dapat menghasilkan temuan dengan tingkat akurasi yang tinggi dan meningkatkan produk akhir. Namun, sebagian besar penelitian sebelumnya mengandalkan penyematan Word2Vec [23] atau TF-IDF [24], yang memiliki keterbatasan dalam menangkap hubungan semantik antar kata [25]. Akibatnya, masih ada sejumlah kesenjangan yang perlu diisi. Selain itu, sebagian besar penelitian sebelumnya berfokus pada platform internasional atau sektor media sosial dan komersial [26], yang berarti hanya sedikit relevansinya dengan konteks pendidikan Indonesia. Melalui analisis peringkat sentimen pengguna pada Platform Merdeka Mengajar (PMM), penelitian ini berupaya menutup kesenjangan dengan mengevaluasi kinerja algoritma RNN dan LSTM berdasarkan embedding FastText [27]. Embedding FastText dipilih dibandingkan TF-IDF atau Word2Vec karena kemampuannya yang unggul dalam menangkap morfologi bahasa dengan menangani istilah di luar kosakata (OOV), yang sering ditemukan dalam tulisan informal bahasa Indonesia [28].

Salah satu hal yang membedakan penelitian ini dari sejumlah penelitian sebelumnya yang telah dijelaskan adalah data yang digunakan. Dengan menggunakan teknik K-Fold Cross Validation [29] dan Confusion Matrix [30], penelitian ini bertujuan untuk meningkatkan akurasi [14] dan stabilitas [31] model dalam memprediksi sentimen positif dan negatif, yang sering diabaikan dalam penelitian sebelumnya [32]. Kedua fitur ini membedakan penelitian ini dari investigasi sebelumnya.

3. METODE PENELITIAN

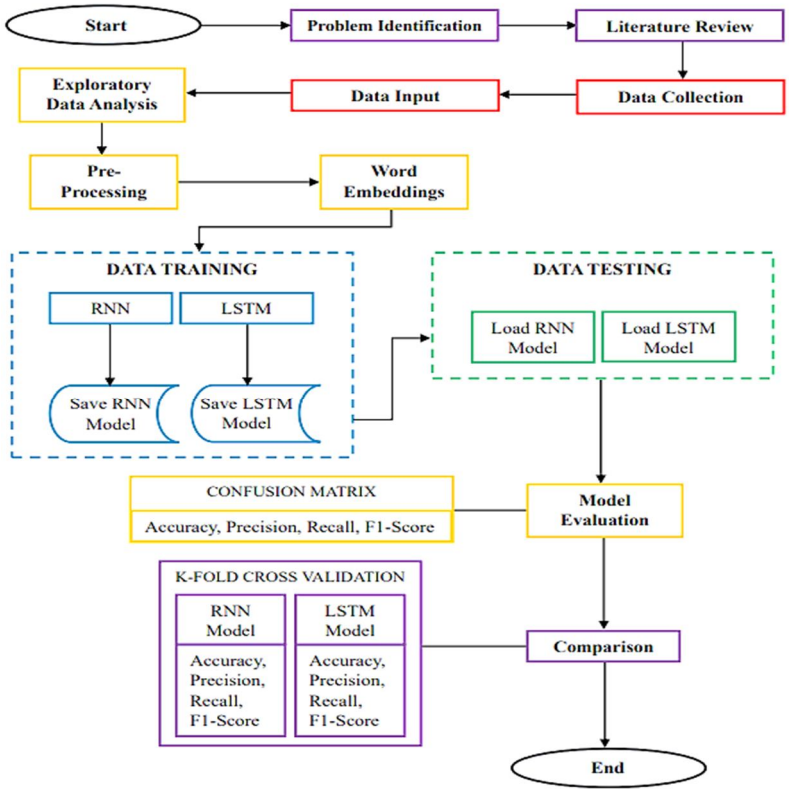
Seperti diilustrasikan pada Gambar 1, penelitian ini menggunakan proses Sampling, Explore, Modify, Model, dan Assess (SEMMA). Model prediktif untuk kumpulan data besar telah berhasil dikembangkan menggunakan metode SEMMA [33]. Pendekatan SEMMA mudah dipahami dan dapat digunakan sebagai panduan untuk proyek penambangan data. Sampel adalah prosedur yang digunakan untuk mengumpulkan informasi dan data penting. Tahap eksplorasi adalah tahap di mana kumpulan data yang berkaitan dengan konsep yang akan dikembangkan dicari. Proses pengelompokan variabel dan pengubahan data disebut "modifikasi." Teknik pemodelan data untuk memprediksi hasil yang

diinginkan disebut model. Akses adalah peninjauan data untuk analisis. Berdasarkan kinerja model, model klasifikasi yang dikembangkan pada langkah sebelumnya akan dievaluasi.



Gambar 1. Metode SEMMA

Memeriksa keluaran model dalam hal recall, akurasi, dan presisi memungkinkan evaluasi. Nilai akurasi menunjukkan seberapa akurat model tersebut, dengan prediksi yang tepat sesuai dengan data yang tersedia. Nilai kerugian, di sisi lain, menunjukkan berapa banyak kesalahan yang ada dalam setiap siklus. Kualitas model yang dihasilkan meningkat seiring dengan penurunan kerugian. Persentase data yang diproyeksikan positif dikenal sebagai presisi. Persentase data positif sejati disebut recall. Dengan 0 sebagai nilai terendah dan 1 sebagai nilai tertinggi, skor F1 adalah rata-rata dari presisi dan recall. [8].



Gambar 2. Diagram Alur Kerja Penelitian

Langkah pertama dalam penelitian ini adalah pengumpulan data, yang melibatkan perolehan data ulasan dari Playstore. Selanjutnya adalah analisis data eksplorasi, yang melihat distribusi frekuensi dan karakteristik data. Terakhir, pra-pemrosesan data, yang melibatkan pembersihan dan transformasi data. Fase berikutnya adalah penyematan kata, yang menggunakan penyematan FastText [13] untuk mengubah data teks menjadi vektor numerik yang dapat digunakan sebagai masukan untuk metode klasifikasi model RNN dan LSTM. Seperti yang terlihat pada Gambar 2, model K-Fold Cross Validation [31] akan menggunakan keluaran dari setiap model dasar sebagai masukan setelah diberikan Confusion Matrix [32].

Pengumpulan Data

Paket Python Scraper digunakan untuk mengumpulkan data ulasan platform Merdeka Mengajar di Google Play Store. Setelah pustaka diinstal melalui pengelola paket, impor pustaka tersebut. Kunci utama untuk pengambilan data adalah ID unik yang diberikan kepada setiap aplikasi Play Store. Setelah menentukan karakteristik bahasa, negara, pengurutan, jumlah ulasan, dan filter, ID yang digunakan dalam penelitian ini adalah id.belajar.app. Terdapat total 17.848 data ulasan, menurut temuan dari pemanfaatan paket pengikis Python Google Colab untuk mengikis ulasan dari Platform Merdeka Mengajar.

Penelitian ini dimulai pada September 2024, dan data dikumpulkan dari Januari 2022 hingga September 2024. Struktur JSON dari data ulasan yang diperoleh mencakup detail seperti ReviewId, userName, userImage, content, score, thumbsUpCount, reviewCreatedVersion, at, replyContent, repliedAt, dan appVersion. Hanya data skor dan konten—yang merupakan ulasan pengguna—yang digunakan dalam penelitian ini. Nilai skor berkisar antara 1 hingga 5, dengan skor 4 atau lebih tinggi menunjukkan sentimen positif dan skor 3 atau lebih rendah menunjukkan sentimen negatif. Skor dan konten kemudian akan disimpan dalam bingkai data.

Pelabelan Data

Informasi yang digunakan dalam penelitian ini terbatas pada konten dan skor yang ditinjau pengguna. Skor tiga atau lebih rendah menunjukkan sentimen negatif, sedangkan skor empat atau lebih tinggi menunjukkan sentimen baik. Skornya berkisar antara 1 hingga 5. Skor dan konten kemudian akan disimpan dalam kerangka data CSV. Kolom 'skor', yang dianggap menggambarkan sentimen, kemudian akan diubah menjadi kolom sentimen dengan kategori 'Positif' atau 'Negatif' sesuai dengan nilai ambang batas 3. Dengan 'konten' sebagai fitur (X) dan sentimen sebagai tujuan (Y), data kemudian dipisahkan menjadi data pelatihan dan data uji, dengan persentase data pelatihan 80% dan proporsi data uji 20%. Kode di bawah ini menunjukkan bagaimana bahasa pemrograman Python digunakan untuk menyelesaikan proses pelabelan data pada tahap ini menggunakan modul pandas di Google Colab.

```
# Membuat kategori skor sentimen (skor: 1-3 negatif, 4-5 positif)
df['sentiment'] = pd.cut(df['score'], bins=[0, 3, 5], labels=['negatif', 'positif'],
include_lowest=True)
# Jika skor > 3 (atau ambang batas lainnya), anggap positif, jika tidak, negatif
df['sentimen'] = df['skor']. terapkan(lambda skor: 'Positif' jika skor > 3, jika tidak 'Negatif')
# Membagi data pelatihan dan pengujian
X = df['konten']
y = df['sentimen']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

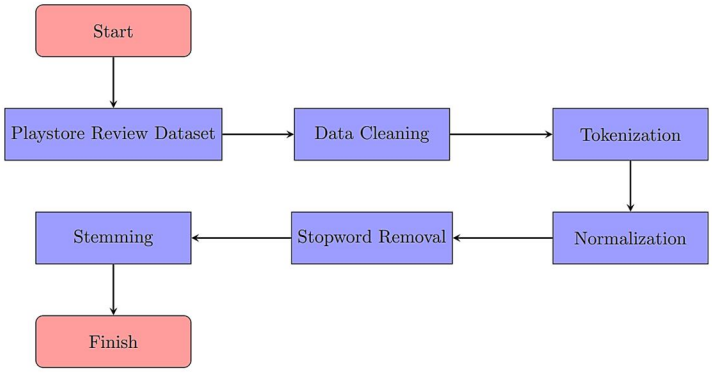
Pra Pemrosesan

Sebelum membuat kategorisasi sentimen pada evaluasi aplikasi Google Play Store, fase pra-pemrosesan sangat penting. Kerangka data harus disiapkan terlebih dahulu untuk diproses dengan menghilangkan kolom yang tidak perlu, hanya menyisakan kolom konten dan skor.

Tabel 1. Data Frame

Content	Score
Jelek selalu tak bener suka error'	1
APLIKASI PMM MEMBUAT BEBAN GURU MAKIN BERAT	1
Kesulitan upload bukti karya	2
Hp iphone tdk bisa download	1
Saya tidak bisa me reset e kin yg telah disetujui.	2

Diagram alir pada Gambar 3 mengilustrasikan bagaimana data ulasan Playstore diproses (Gambar 3), dimulai dengan pengambilan data dan dilanjutkan dengan pembersihan data untuk menghilangkan komponen yang tidak perlu. Setelah itu, data diubah menjadi token, dinormalisasi, kata-kata umum dihilangkan, dan dikembalikan ke bentuk aslinya (stemming) [26]. Perpustakaan Sastrawi Python digunakan dalam penelitian ini [6]. Algoritma stemming Sastrawi memecah kata menjadi bentuk dasarnya dengan menerapkan prinsip morfologi bahasa Indonesia. Data disiapkan untuk digunakan dalam analisis tambahan setelah semua prosedur selesai.



Gambar 3. Tahap Pra Pemrosesan

Proses pengecekan data ulasan untuk nilai yang hilang dan penentuan cara mengelolanya dikenal sebagai pembersihan data. Pada titik ini, data pencilan dihilangkan, data duplikat juga dihapus, dan huruf, tanda baca, simbol khusus seperti "!?#&*", serta simbol-simbol lain yang tidak berarti dihilangkan. Agar model atau algoritma dapat memahami makna setiap kata yang dipisahkan, tokenisasi kemudian melengkapi proses membagi teks ulasan menjadi token atau kata-kata yang berbeda. Semua kata akan diubah menjadi huruf kecil dan dimasukkan ke dalam format yang sama selama fase normalisasi. Selain itu, kata-kata dengan makna yang sama akan digabungkan menjadi satu bentuk.

Eliminasi kata henti adalah proses menghilangkan istilah-istilah populer dalam bahasa Indonesia seperti "dan," "atau," "yang," "ini," dan seterusnya yang tidak memiliki makna. Proses memecah kata menjadi bentuk dasarnya disebut stemming. Perpustakaan Sastrawi Python digunakan dalam penelitian ini. Algoritma stemming Sastrawi memecah kata menjadi bentuk dasarnya menggunakan hukum morfologi bahasa Indonesia. Karena istilah "meminjam" dan "pinjam" berubah menjadi "pinjam" dan "mengangsur" menjadi "angsur," masing-masing, kata-kata yang memiliki makna sama tetapi bentuk berbeda dapat dianggap saling dipertukarkan. Tabel 2 menampilkan hasil praproses data.

Tabel 2. Pra Pemrosesan

Preprocessing	Review	Result
Data Cleaning	Teruntuk para PENELAAH Aksi Nyata, Dengan tegas saya sampaikan, Banyak aksi nyata yg sudah sesuai panduan tetapi di buat perbaikan, sementara saya menemukan ssalah satu akun yg aksi nyatanya cuma 4 slide, tetapi di Validasi. Ini sangat merugikan kami yg membuat aksi nyata dengan nyata beraksi bukan copas lalu upload.	Teruntuk para PENELAAH Aksi Nyata Dengan tegas saya sampaikan Banyak aksi nyata yg sudah sesuai panduan tetapi di buat perbaikan sementara saya menemukan ssalah satu akun yg aksi nyatanya cuma slide tetapi di Validasi Ini sangat merugikan kami yg membuat aksi nyata dengan nyata beraksi bukan copas lalu upload
Tokenization	Teruntuk para PENELAAH Aksi Nyata Dengan tegas saya sampaikan Banyak aksi nyata yg sudah sesuai panduan tetapi di buat perbaikan sementara saya menemukan ssalah satu akun yg aksi nyatanya cuma slide tetapi di Validasi Ini sangat merugikan kami yg membuat aksi nyata dengan nyata beraksi bukan copas lalu upload	["Teruntuk", "para", "PENELAAH", "Aksi", "Nyata", "Dengan", "tegas", "saya", "sampaikan", "Banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "Validasi", "Ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"]
Lowercasing	["Teruntuk", "para", "PENELAAH", "Aksi", "Nyata", "Dengan", "tegas", "saya", "sampaikan", "Banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "Validasi", "Ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"]	["teruntuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"]
Normalization	["teruntuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"]	["untuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yang", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "salah", "satu", "akun", "yang", "aksi", "nyatanya", "hanya", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yang", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copy", "lalu", "unggah"]
Stopword Removal	["untuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yang", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "salah", "satu", "akun", "yang", "aksi", "nyatanya", "hanya", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yang", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copy", "lalu", "unggah"]	["penelaah", "aksi", "nyata", "tegas", "sampaikan", "banyak", "aksi", "nyata", "sesuai", "panduan", "buat", "perbaikan", "menemukan", "salah", "akun", "aksi", "nyatanya", "hanya", "slide", "validasi", "merugikan", "membuat", "aksi", "nyata", "beraksi", "bukan", "copy", "unggah"]

Stemming	["penelaah", "aksi", "nyata", "tegas", "sampaikan", "banyak", "aksi", "nyata", "sesuai", "panduan", "buat", "perbaikan", "menemukan", "salah", "akun", "aksi", "nyatanya", "hanya", "slide", "validasi", "merugikan", "membuat", "aksi", "nyata", "beraksi", "bukan", "copy", "unggah"]	["penelaah", "aksi", "nyata", "tegas", "sampai", "banyak", "aksi", "nyata", "sesuai", "panduan", "buat", "perbaiki", "temu", "salah", "akun", "aksi", "nyata", "hanya", "slide", "validasi", "rugi", "buat", "aksi", "nyata", "beraksi", "bukan", "copy", "unggah"]

Word Embeddings

Penyematan kata dilakukan menggunakan korpus dari FastText Embedding, yang menangani data yang sangat luas dan beragam secara tepat dengan mempelajari penyematan secara dinamis selama pelatihan model. Dalam matriks penyematan, input kata yang di-tokenisasi diubah menjadi vektor numerik menggunakan fungsi penyematan FastText model [28]. X_train dan X_test diubah oleh lapisan embedding menjadi representasi vektor kata berdimensi tetap (128 untuk setiap kata) yang memaksimalkan pelatihan dan menangkap hubungan semantik atau makna kata. Tergantung pada data tertentu yang diberikan, embedding akan mempelajari representasi yang relevan. Facebook AI Research menciptakan teknik penyematan kata yang dikenal sebagai FastText. Dengan belajar mengantisipasi kemungkinan suatu kata diberikan konteks tertentu, konteksnya, atau sebaliknya, FastText mengekspresikan kata-kata sebagai vektor.

Metode FastText memecah kata menjadi n-gram karakter, yang merupakan unit kata yang lebih kecil. FastText sangat membantu untuk bahasa dengan morfologi yang kaya atau saat bekerja dengan kata-kata yang tidak ada dalam leksikon karena dapat menangkap informasi morfologis dan sintaksis dengan mempertimbangkan unit subkata ini. Selama pelatihan, FastText membangun leksikon kata dan unit subkata mereka. Setelah itu, ia belajar memberikan representasi vektor kata dan subkata sesuai dengan pola ko-okurensi dalam korpus teks tertentu. Keterkaitan semantik antara kata dan komponen sub-katanya dapat ditangkap oleh vektor kata yang dihasilkan [27]. Pada tahap Word Embeddings, penelitian ini menggunakan FastText Embeddings dengan korpus berikut:

[['mantap'], ['sangat', 'membantu'], ['mantap'], ['bagus'], ['sangat', 'baik', 'dan', 'memuaskan', 'dan', 'yg', 'utama', 'sangat', 'membantu', 'dalam', 'mengajar'], ['sangat', 'mempermudah'], ['mantab'], ['bagus'], ['sangat', 'membantu', 'sebagai', 'bahan', 'referensi'], ['membantu', 'guru', 'hebat']]

Menggunakan embedding FastText, yang digunakan dalam penelitian ini, embedding kata membuat analisis teks lebih efisien dengan merepresentasikan kata sebagai vektor berdimensi tinggi yang menangkap hubungan semantik antar kata dan morfologi bahasa.

Pemodelan

Klasifikasi adalah tindakan selanjutnya yang harus diambil. Dalam analisis data, klasifikasi menjadi langkah penting yang membantu mengorganisasi data ke dalam kelas atau kategori berdasarkan ciri atau atribut tertentu. Klasifikasi membantu dalam menemukan pola atau ciri yang relevan dalam suatu kumpulan data, yang memungkinkan kita untuk memprediksi kelas data baru menggunakan model yang telah dilatih dalam konteks penambangan data dan pembelajaran mesin [34]. Pendekatan ini sangat membantu dalam berbagai bidang, termasuk sistem rekomendasi untuk mengklasifikasikan produk yang menarik bagi pengguna, diagnostik medis untuk mengklasifikasikan pasien berdasarkan gangguan, dan analisis sentimen untuk mengklasifikasikan evaluasi positif atau negatif.

Pada titik ini, dikembangkan model analisis sentimen untuk menilai ulasan pengguna terhadap platform Merdeka Mengajar. Karena sifatnya yang berurutan dan arsitektur yang disesuaikan untuk data berbentuk urutan dan daftar, Jaringan Saraf Berulang (RNN) dipilih sebagai model klasifikasi [12]. RNN memproses input berurutan dan menggunakan metode loop internal untuk mempertahankan

informasi dari langkah-langkah sebelumnya. Model ideal juga dikembangkan dalam prosedur ini menggunakan LSTM [35]. Instrumen untuk memeriksa teks evaluasi yang sudah ada adalah LSTM. Gerbang masukan, yang memilih nilai mana yang akan diperbarui; gerbang lupakan, yang harus memutuskan informasi mana yang akan digunakan atau tidak; dan gerbang keluaran, yang memilih konteks yang akan dihasilkan, adalah tiga tahap pertama dalam proses komputasi LSTM. Dengan mengontrol aliran informasi menggunakan gerbang unik, LSTM memungkinkan model untuk mempertahankan dan mengingat data untuk jangka waktu yang lebih lama.

Evaluasi

Setelah pemodelan klasifikasi selesai pada langkah sebelumnya, penilaian kinerja pemodelan klasifikasi pembelajaran mesin dilakukan pada langkah ini. Tujuan dari langkah ini adalah untuk mengevaluasi kinerja dan efektivitas kedua model klasifikasi pembelajaran mesin. Matriks kebingungan adalah metode yang digunakan pada titik ini. Hasil keseluruhan dari klasifikasi yang benar dan salah diringkas menggunakan matriks kebingungan. Kombinasi data True Positive (TP), False Negative (FN), False Positive (FP), dan True Negative (TN) dapat digunakan untuk melihat matriks kebingungan. Menggunakan rumus dalam Tabel 3, matriks kebingungan menghasilkan empat hasil pengukuran: akurasi, presisi, recall, dan skor F1 [30].

Tabel 3. Struktur *Confusion Matrix*

	Positive Prediction	Negative Prediction
Positive Actual	True Positive (TP)	False Negative (FN)
Negative Actual	False Positive (FP)	True Negative (TN)

Pendekatan K-Fold Cross Validation digunakan dalam langkah evaluasi selanjutnya. Dengan nilai k=5, metode ini menghasilkan lima hasil yang ditampilkan untuk metrik akurasi, presisi, recall, dan skor F1 yang diperoleh dari prosedur Validasi Silang K-Fold. Data dibagi menjadi lima bagian (fold) untuk Validasi Silang K-Fold, dengan satu bagian berfungsi sebagai data uji dan empat bagian lainnya sebagai data pelatihan untuk setiap iterasi. Pada tahun 2022, Azizah dkk. menghasilkan lima set metrik penilaian untuk setiap model yang diuji sebagai hasil dari proses ini yang dilakukan lima kali [31]. Hal ini dilakukan beberapa kali untuk mendapatkan perkiraan kinerja yang lebih andal dan presisi.

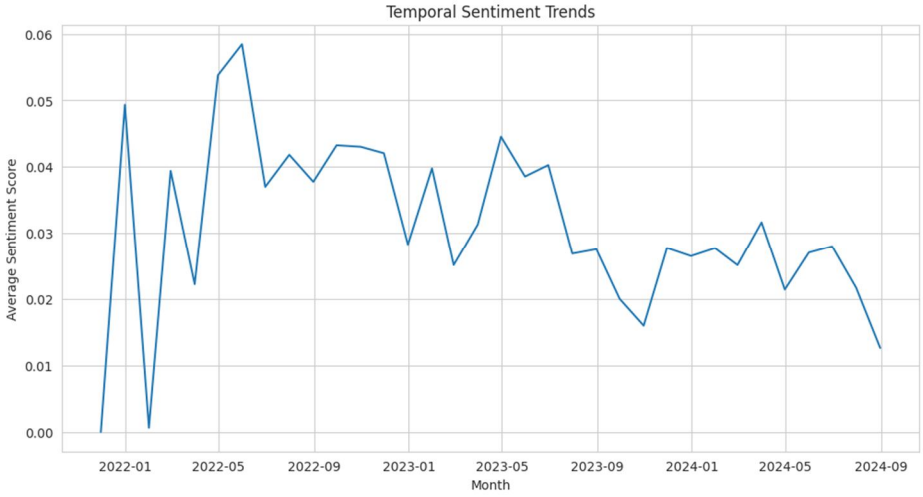
4. HASIL DAN PEMBAHASAN

Distribusi frekuensi skor ulasan aplikasi Platform Merdeka Mengajar ditampilkan pada Tabel 4, di mana skor 5 dengan frekuensi lebih dari 14.000 ulasan mendominasi, menunjukkan kepuasan pelanggan yang sangat tinggi. Sebaliknya, skor 4 menunjukkan tingkat sedang, sedangkan skor rendah seperti 1, 2, dan 3 memiliki frekuensi yang jauh lebih rendah.

Tabel 4. Distribusi Ulasan Pengguna

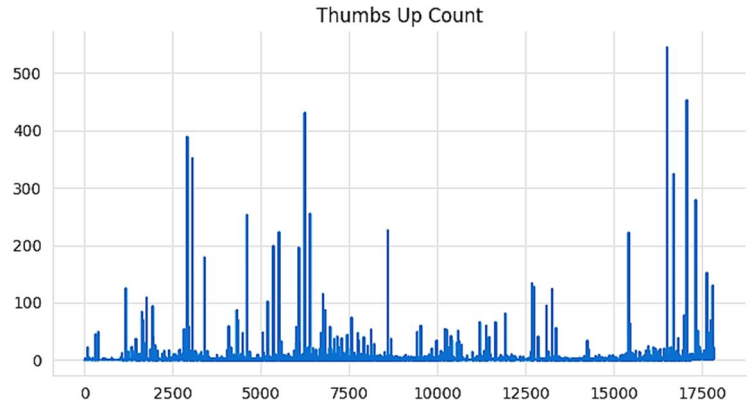
Scores	Frequency	Percentage (%)
1	712	3,9892
2	207	1,1598
3	497	2,7846
4	1503	8,4211
5	14929	83,6452
Total	17,848	100,00

Kemudian dihitung distribusi kelas sentimen. Data yang ditunjukkan pada Tabel 4 mengindikasikan bahwa kelas sentimen positif adalah 16432 (92,1%), sedangkan kelas sentimen negatif adalah 1416 (7,9%). Ulasan positif menunjukkan pengalaman pengguna Merdeka Mengajar yang positif dan memuaskan. Ulasan negatif mengungkapkan masalah atau ketidakpuasan yang dialami pelanggan, seperti kesulitan menggunakan program, fitur yang hilang, atau bug yang mengganggu.



Gambar 4. Tren Sentimen Temporal

Grafik temporal pada Gambar 4 dengan demikian menunjukkan tren peningkatan dalam peringkat positif, tetapi beberapa puncak juga menunjukkan lonjakan ulasan negatif, seperti ketika pembaruan aplikasi tertentu dirilis. Ulasan negatif mengungkap istilah-istilah seperti "bug" dan "akses sulit," yang menggambarkan masalah spesifik yang dihadapi pelanggan. Distribusi sentimen ulasan juga menunjukkan jumlah perasaan positif yang signifikan. Menyajikan data ini secara lebih rinci selama debat dapat menghasilkan wawasan yang lebih mendalam tentang tren pengalaman pengguna.

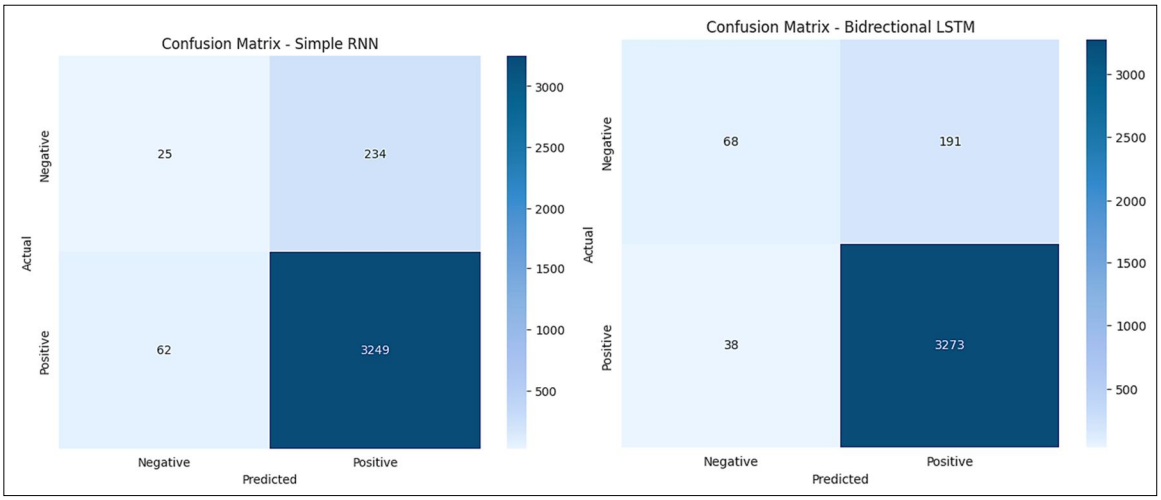


Gambar 5. Plot Data Hitungan *Like*

Gambar 5 memplot indeks data (baris data) pada sumbu horizontal (X) terhadap data ThumbsUpCount (jumlah suka) pada sumbu vertikal (Y). Banyak nilai data rendah (sekitar 0), sementara beberapa indeks memiliki puncak yang signifikan (nilai tinggi). Puncak tertinggi, yang menunjukkan bahwa data tertentu memiliki jumlah suka yang jauh lebih tinggi daripada data lainnya, terletak dekat dengan indeks 17500.

Hasil

Data dari *FastText Embeddings*, yang sudah dalam bentuk vektor numerik, digunakan sebagai input untuk model algoritma dengan persentase data pelatihan 80% dan persentase data pengujian 20% pada tahap perbandingan matriks kebingungan antara algoritma RNN dan LSTM.



Gambar 5. Confusion Matrix – Perbandingan RNN & LSTM

Setiap model RNN dan LSTM menjalani analisis sentimen selama periode 20 epoch. Tabel matriks kebingungan dengan peta panas berwarna yang membandingkan nilai yang diharapkan dan nilai sebenarnya kemudian digunakan untuk menampilkan hasilnya.

Hasil pengujian pada Tabel 5 menunjukkan bahwa model RNN mencapai akurasi 91,70%, presisi 88,59%, recall 91,70%, dan skor F1 89,75%. Sementara itu, algoritma LSTM mencapai akurasi 93,58%, presisi 92,28%, recall 93,58%, dan skor F1 92,23%. Hasil ini mendukung hipotesis bahwa algoritma berbasis LSTM lebih unggul dalam menangkap pola kompleks dibandingkan RNN untuk data berurutan ulasan pengguna pada platform pendidikan.

Tabel 5. Perbandingan Hasil Akurasi RNN & LSTM

Algorithm	Accuracy	Precision	Recall	F1-Score
RNN	**91.70%**	88.59%	91.70%	89.75%
LSTM	**93.58%**	92.28%	93.58%,	92.23%.

Hasil proses validasi silang K-Fold Cross Validation ditunjukkan pada Tabel 6, di mana akurasi rata-rata, presisi, recall, dan skor F1 model RNN masing-masing adalah 92,23%, 89,06%, 92,23%, dan 89,12% saat menggunakan model RNN dan LSTM. Model tersebut berkinerja lebih baik dengan LSTM, mencapai akurasi rata-rata 92,70%, presisi 91,03%, recall 92,70%, dan skor F1 90,68%.

Tabel 6. Perbandingan Rata-rata Akurasi RNN & LSTM

Algorithm	Accuracy	Precision	Recall	F1-Score
RNN	**92,23%**	89,06%	92,23%	89,12%
LSTM	**92,70%**	91,03%	92,70%	90,68%

Berdasarkan temuan ini, LSTM terus mengungguli RNN di semua metrik evaluasi. Durasi setiap langkah dalam proses pelatihan kedua model tersebut juga menunjukkan bahwa LSTM membutuhkan

waktu sedikit lebih lama dibandingkan RNN, yang masuk akal mengingat arsitektur LSTM lebih rumit karena mekanisme gerbang lupakan dan keadaan sel dalam memproses data berurutan. Karena teknik K-Fold Cross Validation dapat mengurangi bias dalam pembagian data, evaluasi model menjadi lebih stabil.

Tabel 7. Perbandingan Penelitian Terdahulu

Penelitian	Algoritma	*Embeddings*	Ruang Lingkup	Dataset	Akurasi Terbaik
Nemes & Kiss (2021)	RNN	Word2Vec	COVID-19 on Twitter	500	90%
Widayat (2021)	LSTM	Word2Vec	Movie Review	25,000	88.17%
Isnain et al. (2022)	LSTM & Naive Bayes	TF-IDF	Kebijakan *New Normal* di Twitter	15,000	83.33% (LSTM)
Utami (2022)	RNN	TF-IDF	Shopee application (Online Shop)	800	80%
Subowo et al. (2022)	LSTM	FastText	E-commerce	1000	81%
Penelitian ini	RNN & LSTM	FastText	Educational Platforms	17,848	**91.70%** (RNN), **93.58%** (LSTM)

Temuan penelitian ini dibandingkan dengan penelitian analisis sentimen sebelumnya yang menggunakan algoritma RNN dan LSTM dengan strategi dan pengaturan embedding yang bervariasi dalam Tabel 7. Dibandingkan dengan penelitian sebelumnya seperti Subowo et al. (2022)[22] dengan akurasi 81% dalam konteks e-commerce dan Widayat (2021)[21] dengan 88,17% dalam ulasan film, hasil penelitian ini menunjukkan akurasi tertinggi sebesar 93,58% menggunakan LSTM dengan embedding FastText dalam konteks platform pendidikan. Kemampuan FastText dalam menangani kata-kata di luar kosakata (OOV) dan kerumitan bahasa Indonesia, yang belum diteliti secara mendalam dalam penelitian sebelumnya, memungkinkan LSTM mengungguli model lain dalam penelitian ini.

Dengan demikian, penelitian ini menunjukkan seberapa baik LSTM dan FastText bekerja untuk meningkatkan akurasi model dan membuat kemajuan signifikan dalam analisis sentimen, khususnya di bidang platform pendidikan.

Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Long Short-Term Memory (LSTM) memiliki keunggulan signifikan dibandingkan Recurrent Neural Network (RNN) dalam analisis sentimen ulasan pengguna pada platform Merdeka Mengajar. Dengan akurasi 93,58%, presisi 92,28%, recall 93,58%, dan skor F1 92,23%, LSTM terbukti lebih efektif dalam menangkap hubungan temporal yang kompleks melalui mekanisme gerbang lupakan dan keadaan sel. Di sisi lain, meskipun RNN mencapai akurasi tinggi sebesar 91,70%, arsitektur ini memiliki keterbatasan dalam menangani konteks jangka panjang. Perbedaan ini konsisten dengan temuan Isnain et al. (2022)[19], yang menunjukkan bahwa LSTM unggul dalam menangkap morfologi bahasa dan konteks semantik dibandingkan dengan RNN. Namun, LSTM membutuhkan waktu pelatihan yang lebih lama, yang menjadi tantangan bagi aplikasi waktu nyata.

Validasi silang dengan teknik K-Fold mengonfirmasi bahwa LSTM lebih stabil dibandingkan dengan RNN, dengan peningkatan akurasi rata-rata sebesar 1,47% selama validasi. Temuan ini menunjukkan keandalan LSTM dalam generalisasi data. Analisis distribusi sentimen mengungkapkan bahwa mayoritas ulasan pengguna positif (92,1%), mencerminkan pengalaman yang baik. Namun, ulasan negatif (7,9%) mencatat keluhan seperti bug aplikasi dan keterbatasan fitur, yang mendukung temuan Ketaren et al. (2022)[4] tentang tantangan teknis dalam platform pendidikan.

Penerapan embedding FastText menunjukkan efektivitas teknik ini dalam menangkap hubungan semantik antar kata, terutama dalam bahasa Indonesia yang kompleks. Representasi berbasis FastText berkontribusi signifikan terhadap peningkatan akurasi model, terutama dalam menangani kata-kata di luar kosakata. Hal ini mendukung penelitian Ariyus & Manongga (2024)[27], meskipun diperlukan studi lebih lanjut mengenai pengaruh parameter embedding untuk dataset dengan konteks lokal. Penelitian ini memiliki kekuatan utama dalam penggunaan embedding FastText dan validasi silang yang andal, tetapi juga memiliki keterbatasan. Ketidakseimbangan dalam distribusi sentimen dapat menyebabkan bias, jadi teknik oversampling atau metrik evaluasi tambahan seperti Koefisien Korelasi Matthews perlu dipertimbangkan. Selain itu, kumpulan data yang hanya mencakup ulasan Google Play Store mungkin tidak sepenuhnya mencerminkan pengalaman guru di berbagai wilayah.

Hasil ini memiliki implikasi penting bagi pengembang aplikasi, seperti memperbaiki bug, meningkatkan fitur, dan mengoptimalkan aksesibilitas aplikasi untuk mendukung pengguna di area dengan keterbatasan jaringan. Studi di masa depan direkomendasikan untuk mengeksplorasi arsitektur model yang lebih kompleks, seperti model berbasis *transformer*, dan untuk memperluas dataset agar mencakup survei lapangan guna memberikan wawasan yang lebih komprehensif.

5. KESIMPULAN

Penelitian ini membahas bagaimana algoritma pembelajaran mendalam seperti LSTM dan RNN dapat diterapkan untuk meningkatkan akurasi analisis sentimen pada ulasan pengguna platform pendidikan digital Merdeka Mengajar. Keunggulan LSTM dalam menangkap hubungan temporal melalui mekanisme gerbang lupakan dan keadaan sel memberikan keunggulan signifikan untuk menganalisis data sekuensial yang kompleks. Distribusi ulasan yang sebagian besar positif (92,1%) juga mencerminkan tingkat kepuasan pengguna yang tinggi terhadap fitur platform, khususnya dalam aspek pelatihan mandiri. Implementasi embedding FastText dalam penelitian ini terbukti efektif dalam menangkap hubungan semantik antar kata, terutama untuk bahasa Indonesia yang kompleks. Hal ini berkontribusi pada peningkatan akurasi klasifikasi sentimen, terutama dalam menangani kata-kata di luar kosakata yang sering ditemukan dalam ulasan informal.

Dengan akurasi 93,58%, LSTM mengungguli model RNN yang memiliki akurasi 91,70%. Namun, penelitian ini memiliki keterbatasan, termasuk bias data dari distribusi sentimen yang tidak merata (92,1% positif) dan ketergantungan pada ulasan Google Play Store, yang mungkin tidak secara akurat mencerminkan pengalaman pengguna di Indonesia. 7,9% evaluasi negatif mencakup berbagai keluhan tentang keterbatasan fitur, gangguan aplikasi, dan masalah akses di lokasi terpencil. Menurut analisis ini, pengembang harus memprioritaskan pengujian aplikasi mereka di berbagai perangkat dan kondisi jaringan. Mereka juga harus mempertimbangkan untuk menambahkan fitur seperti mode offline untuk mendukung aksesibilitas di tempat-tempat dengan infrastruktur yang tidak memadai. Wawasan ini akan membantu pengembang aplikasi pendidikan meningkatkan kualitas layanan dan fitur platform mereka.

DAFTAR PUSTAKA

[1] M. G. Maru, F. P. Tamowangkay, N. Pelenkahu, and C. Wuntu, ‘Teachers’ perception toward the impact of platform used in online learning communication in the eastern Indonesia’, *Int. J. Commun. Soc.*, vol. 4, no. 1, pp. 59–71, 2021, doi: 10.31763/ijcs.v4i1.321.

[2] W. Ndari, Suyatno, Sukirman, and F. N. Mahmudah, ‘Implementation of the Merdeka Curriculum and Its Challenges’, *Eur. J. Educ. Pedagog.*, vol. 4, no. 3, pp. 111–116, 2023, doi: 10.24018/ejedu.2023.4.3.648.

[3] T. T. D. S. Iqmatul Pratiwi, Desi Rahmawati, ‘Lectura : Jurnal Pendidikan’, vol. 15, pp. 596–610, 2024.

[4] A. Ketaren, F. Rahman, H. P. Meliala, N. Tarigan, and R. Simanjuntak, 'Monitoring dan Evaluasi Pemanfaatan Platform Merdeka Mengajar pada Satuan Pendidikan Aswinta', *J. Pendidik. dan Konseling*, vol. 4, no. 6, pp. 10340–10343, 2022, doi: https://doi.org/10.31004/jpdk.v4i6.10030.

[5] H. Gomez-Adorno, G. Bel-Enguix, G. Sierra, J. C. Barajas, and W. Álvarez, 'Machine Learning and Deep Learning Sentiment Analysis Models: Case Study on the SENT-COVID Corpus of Tweets in Mexican Spanish', *Informatics*, vol. 11, no. 2, 2024, doi: 10.3390/informatics11020024.

[6] A. Munna and E. Zuliarso, 'Interpretasi model Stacking Ensemble untuk analisis sentimen ulasan aplikasi pinjaman online menggunakan LIME', vol. 21, no. 2, pp. 183–196, 2024.

[7] M. S. David and S. Renjith, 'Comparison of word embeddings in text classification based on RNN and CNN', *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1187, no. 1, p. 012029, 2021, doi: 10.1088/1757-899x/1187/1/012029.

[8] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa'ah, 'Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes', *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 802–808, 2021, doi: 10.29207/resti.v5i4.3308.

[9] L. Kurniasari and A. Setyanto, 'Sentiment Analysis using Recurrent Neural Network', *J. Phys. Conf. Ser.*, vol. 1471, no. 1, 2020, doi: 10.1088/1742-6596/1471/1/012018.

[10] L. Khan, A. Amjad, K. M. Afaq, and H. T. Chang, 'Deep Sentiment Analysis Using CNN-LSTM Architecture of English and Roman Urdu Text Shared in Social Media', *Appl. Sci.*, vol. 12, no. 5, 2022, doi: 10.3390/app12052694.

[11] L. Atikah, D. Purwitasari, and N. Suciati, 'DETEKSI KEJADIAN LALU LINTAS PADA TEKS TWITTER DENGAN PENDEKATAN KLASIFIKASI MULTI-LABEL BERBASIS DEEP LEARNING MULTI-LABEL CLASSIFICATION USING DEEP LEARNING APPROACH ON TWITTER', vol. 9, no. 1, pp. 87–96, 2022, doi: 10.25126/jtiik.202295206.

[12] T. A. Zuraiyah, M. M. Mulyati, and G. H. F. Harahap, 'Perbandingan Metode Naïve Bayes, Support Vector Machine Dan Recurrent Neural Network Pada Analisis Sentimen Ulasan Produk E-Commerce', *Multitek Indones.*, vol. 17, no. 1, pp. 27–43, 2023, doi: 10.24269/mtkind.v17i1.7092.

[13] A. S. A. Al-jumaili, 'NAMED ENTITY RECOGNITION', vol. 25, pp. 5258–5264, 2024, doi: 10.12694/scpe.v25i6.3365.

[14] C. Liu *et al.*, 'Improving sentiment analysis accuracy with emoji embedding', *J. Saf. Sci. Resil.*, vol. 2, no. 4, pp. 246–252, 2021, doi: 10.1016/j.jnlssr.2021.10.003.

[15] R. Maulana, P. A. Rahayuningsih, W. Irmayani, D. Saputra, and W. E. Jayanti, 'Improved Accuracy of Sentiment Analysis Movie Review Using Support Vector Machine Based Information Gain', *J. Phys. Conf. Ser.*, vol. 1641, no. 1, 2020, doi: 10.1088/1742-6596/1641/1/012060.

[16] S. Shumaly, M. Yazdinejad, and Y. Guo, 'Persian sentiment analysis of an online store independent of pre-processing using convolutional neural network with fastText embeddings', *PeerJ Comput. Sci.*, vol. 7, pp. 1–22, 2021, doi: 10.7717/peerj-cs.422.

[17] K. L. Tan, C. P. Lee, K. S. M. Anbananthen, and K. M. Lim, 'RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network', *IEEE Access*, vol. 10, pp. 21517–21525, 2022, doi: 10.1109/ACCESS.2022.3152828.

[18] L. Nemes and A. Kiss, 'Social media sentiment analysis based on COVID-19', *J. Inf. Telecommun.*, vol. 5, no. 1, pp. 1–15, 2021, doi: 10.1080/24751839.2020.1790793.

[19] A. R. Isnain, H. Sulistiani, B. M. Hurohman, A. Nurkholis, and S. Styawati, 'Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen', *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 2, p. 299, 2022, doi: 10.26418/jp.v8i2.54704.

[20] H. Utami, 'Analisis Sentimen dari Aplikasi Shopee Indonesia Menggunakan Metode Recurrent Neural Network', *Indones. J. Appl. Stat.*, vol. 5, no. 1, p. 31, 2022, doi:

10.13057/ijas.v5i1.56825.

[21] W. Widayat, 'Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning', *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 1018, 2021, doi: 10.30865/mib.v5i3.3111.

[22] E. Subowo, F. Adi Artanto, I. Putri, and W. Umaedi, 'Algoritma Bidirectional Long Short Term Memory untuk Analisis Sentimen Berbasis Aspek pada Aplikasi Belanja Online dengan Cicilan', *J. Fasilkom*, vol. 12, no. 2, pp. 132–140, 2022, doi: 10.37859/jf.v12i2.3759.

[23] D. E. Cahyani and I. Patasik, 'Performance comparison of tf-idf and word2vec models for emotion text classification', *Bull. Electr. Eng. Informatics*, vol. 10, no. 5, pp. 2780–2788, 2021, doi: 10.11591/eei.v10i5.3157.

[24] H. D. Abubakar and M. Umar, 'Sentiment Classification: Review of Text Vectorization Methods: Bag of Words, Tf-Idf, Word2vec and Doc2vec', *SLU J. Sci. Technol.*, vol. 4, no. 1&2, pp. 27–33, 2022, doi: 10.56471/slujst.v4i.266.

[25] 2022 Dang, N. C., Moreno-García, M. N., & De la Prieta, 'Sentiment Analysis Based on Deep Learning in E-Commerce', *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13369 LNAI, pp. 498–507, 2022, doi: 10.1007/978-3-031-10986-7_40.

[26] Nadia Ristya Dewi, E. Yulia Puspaningrum, and H. Maulana, 'Analisis Sentimen Tweet Vaksinasi Covid-19 Menggunakan RNN Dengan Metode TF-IDF Dan Word2Vec', *J. Inform. dan Sist. Inf.*, vol. 3, no. 1, pp. 56–65, 2022, doi: 10.33005/jifosi.v3i1.449.

[27] D. Ariyus and D. Manongga, 'Enhancing Sentiment Analysis of Indonesian Tourism Video Content Commentary on TikTok : A FastText and Bi-LSTM Approach', vol. 14, no. 6, pp. 18020–18028, 2024.

[28] Darussalam and G. Arief, 'Jurnal Resti', *Resti*, vol. 1, no. 1, pp. 19–25, 2021.

[29] K. P. C. Kalaiarasan, '"Web Services Performance Prediction with Confusion Matrix and K-Fold Cross Validation to Provide Prior Service Quality Characteristics"', *J. Electr. Syst.*, vol. 20, no. 2s, pp. 284–292, 2024, doi: 10.52783/jes.1139.

[30] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, 'Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking', *IEEE Access*, vol. 8, pp. 90847–90861, 2020, doi: 10.1109/ACCESS.2020.2994222.

[31] R. A. Azizah, F. Bachtiar, and S. Adinugroho, 'Klasifikasi Kinerja Akademik Siswa Menggunakan Neighbor Weighted K-Nearest Neighbor dengan Seleksi Fitur Information Gain', *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, pp. 605–614, 2022, doi: 10.25126/jtiik.2022935751.

[32] R. Sharma, S. Shrivastava, S. Kumar Singh, A. Kumar, S. Saxena, and R. Kumar Singh, 'Deep-Abppred: Identifying antibacterial peptides in protein sequences using bidirectional LSTM with word2vec', *Brief. Bioinform.*, vol. 22, no. 5, pp. 1–19, 2021, doi: 10.1093/bib/bbab065.

[33] Y. A. Suwitono and F. J. Kaunang, 'Implementasi Algoritma Convolutional Neural Network (CNN) Untuk Klasifikasi Daun Dengan Metode Data Mining SEMMA Menggunakan Keras', *J. Komtika (Komputasi dan Inform.*, vol. 6, no. 2, pp. 109–121, 2022, doi: 10.31603/komtika.v6i2.8054.

[34] Z. Darojah, R. Susetyoko, and N. Ramadijanti, 'Strategi Penanganan Imbalance Class Pada Model Klasifikasi Penerima Kartu Indonesia Pintar Kuliah Berbasis Neural Network Menggunakan Kombinasi SMOTE dan ENN', *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 2, pp. 457–466, 2023, doi: 10.25126/jtiik.20231026480.

[35] M. A. Amrustian, W. Widayat, and A. M. Wirawan, 'Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM', *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 535, 2022, doi: 10.30865/mib.v6i1.3527.