

Sentimen analisis Berbasis Aspek pada Ulasan Pengguna aplikasi Sekolah.mu di Google Play store Menggunakan Algoritma *Convolutional Neural Network* (CNN)

1st Jesika Damayanti Doloksaribu
Telkom University
Bandung, Indonesia
jesikadoloksaribu@student.telkomuniversity.ac.id

2nd Oktariani Nurul Pratiwi
Telkom University
Bandung, Indonesia
onurulp@telkomuniversity.ac.id

3rd Nur Ichsan Utama,
Telkom University
Bandung, Indonesia
nichsan@telkomuniversity.ac.id

Perkembangan teknologi digital dan platform pembelajaran daring di Indonesia, seperti Sekolah.mu, telah mengubah proses belajar mengajar. Dengan semakin banyaknya pengguna, muncul ulasan tidak terstruktur di Google Play Store yang dapat menjadi wawasan berharga bagi pengembang aplikasi. Untuk menguraikan data ulasan, diperlukan analisis sentimen berbasis aspek (ABSA) untuk memberikan informasi yang lebih terperinci. Penelitian ini menerapkan ABSA menggunakan algoritma Convolutional Neural Network (CNN) untuk mengklasifikasikan sentimen dan aspek dari ulasan aplikasi Sekolah.mu. Data dari ulasan diproses melalui tahapan Knowledge Discovery in Databases (KDD) dan direpresentasikan dalam vektor oleh Word2Vec. Hasil penelitian menunjukkan bahwa model CNN berhasil mengklasifikasikan sentimen dengan akurasi antara 82% dan 89% di tiga dimensi fokus: konten pendidikan, kinerja aplikasi, dan pengalaman pengguna termasuk metrik keterlibatan pelajar. Pekerjaan ini juga menyediakan aplikasi bertenaga Streamlit yang dikembangkan untuk membantu pengguna dalam memvisualisasikan hasil analisis sentimen yang dilakukan pada ulasan pelajar. Temuan ini menunjukkan bahwa CNN berguna untuk menganalisis konten yang dibuat pengguna secara sengaja dan dapat secara efektif meningkatkan pengalaman pengguna dalam aplikasi pendidikan.

Kata Kunci—Convolutional Neural Network, Aspect-Based Sentiment Analysis, Word2Vec, Streamlit, Aplikasi Sekolah.mu.

PENDAHULUAN

Penggunaan teknologi digital yang semakin meningkat telah membawa pengaruh yang sangat signifikan dalam berbagai aspek kehidupan, tanpa terkecuali pendidikan. Aplikasi Sekolah.mu salah satu platform pembelajaran daring yang dirancang khusus untuk mendukung proses belajar-mengajar dengan berbagai fitur yang terus diperbarui. Sekolah.mu terus berupaya melakukan pengembangan agar sesuai dengan kebutuhan pengguna dan mampu memberikan pengalaman belajar yang lebih baik [1].

Saat ini, Sekolah.mu sedang dalam tahap pengembangan aplikasi untuk meningkatkan kualitas layanan pembelajaran daring salah satunya melalui Google Play

Store. Google Play Store menyediakan platform ulasan dan rating bagi pengguna untuk mengutarakan opini mereka [2]. Ulasan-ulasan tersebut dapat menjadi masukan yang sangat berharga bagi pengembang aplikasi, namun pengolahan ulasan secara manual menghadapi berbagai kendala, seperti tingginya volume ulasan, keberagaman gaya bahasa pengguna, serta redundansi informasi yang menyulitkan penyaringan ulasan yang relevan. Masalah-masalah ini menunjukkan perlunya pendekatan yang lebih efisien seperti sentimen analisis [3].

Salah satu pendekatan yang dapat digunakan dalam memahami opini pengguna adalah melalui sentimen analisis berbasis aspek (Aspect-Based Sentiment Analysis/ABSA). Metode ini telah banyak digunakan dalam berbagai penelitian untuk memahami sentimen pengguna pada berbagai aplikasi populer, namun belum ada penelitian yang secara khusus mengevaluasi ulasan pengguna pada aplikasi Sekolah.mu [4]. ABSA (Sentimen analisis Berbasis Aspek) memungkinkan penyajian opini pengguna bukan hanya pada sentimen keseluruhan tetapi secara mendalam dan terfokus pada aspek-aspek yang lebih spesifik dalam aplikasi [5].

Dalam sentimen analisis berbasis aspek, penelitian terbaru menunjukkan bahwa algoritma Convolutional Neural Network (CNN) jauh lebih efektif untuk analisis teks dibandingkan dengan algoritma lain seperti SVM ataupun LSTM [6]. Namun, pada penelitian terbaru menunjukkan bahwa algoritma Convolutional Neural Network (CNN) jauh lebih efektif untuk analisis teks, khususnya mengekstraksi pola fitur yang terdapat dalam data ulasan pengguna. CNN memiliki kemampuan dalam mengenali pola spasial pada teks yang menjadikannya pilihan yang tepat untuk diterapkan dalam ABSA [7].

Melalui pendekatan ini, hasil penelitian diharapkan dapat memberikan gambaran yang lebih komprehensif mengenai sentimen pengguna, sehingga dapat membantu pengembang aplikasi dalam mengidentifikasi dan melakukan peningkatan untuk memenuhi kebutuhan pengguna dengan lebih baik.

I. KAJIAN TEORI

A. *Natural Language Processing* (NLP)

Natural Language Processing (NLP) merupakan cabang kecerdasan buatan yang dibuat untuk mengeksplorasi hubungan yang rumit antara komputer dengan bahasa manusia. Tujuan utama NLP adalah memahami, memeriksa, dan menghasilkan informasi tekstual dengan cara yang meniru kemampuan manusia [8].

B. *Knowledge Discovery in Database* (KDD)

Knowledge Discovery in Database (KDD) adalah proses untuk mengekstraksi pengetahuan yang berguna dari kumpulan data yang kompleks. Proses ini melibatkan beberapa tahapan dimulai dari pengumpulan data, pembersihan data, transformasi data, pemilihan aspek yang relevan, juga penerapan algoritma untuk menemukan pola atau informasi yang memiliki makna [9].

C. Sentimen Analisis Berbasis Aspek

Sentimen Analisis Berbasis Aspek (*Aspect-Based Sentiment Analysis* atau *ABSA*) merupakan teknik dalam sentimen analisis yang memiliki tujuan untuk mengekstraksi dan menilai sentimen terhadap aspek-aspek tertentu dalam sebuah teks ulasan pengguna terhadap suatu produk atau layanan. Berbeda dengan sentimen analisis umum yang hanya mengukur sentimen secara keseluruhan, *ABSA* memecah ulasan menjadi berbagai elemen atau aspek yang relevan dan mengsentimen analisis untuk setiap elemen tersebut [10].

D. *Preprocessing* data

Preprocessing data merupakan salah satu tahap penting dalam analisis data dan data mining yang memiliki tujuan untuk mempersiapkan data mentah agar lebih siap dan optimal untuk analisis atau pemodelan lebih lanjut. Tahap ini mencakup beberapa langkah utama, seperti pembersihan data, transformasi, seleksi aspek, encoding variabel kategorikal, dan pengurangan dimensi [11].

E. *Split* data

Split data membagi dataset menjadi beberapa bagian agar model dapat dilatih dan diuji. Proses ini biasanya membagi data menjadi dua set utama: data pelatihan, yang digunakan untuk mengajarkan model, memungkinkan model untuk belajar dari pola yang ada dalam data, dan data pengujian, yang digunakan untuk mengevaluasi sejauh mana model yang telah dilatih dapat menggeneralisasi dan bekerja pada data yang belum pernah dilihat sebelumnya, yang sangat penting untuk menghindari *overfitting* [12].

F. *Word2Vec*

Word2Vec merupakan algoritma pembelajaran mendalam yang digunakan untuk mempelajari representasi vektor dalam kata dari data teks. Teknik ini digunakan untuk menghasilkan *embedding* kata, yaitu representasi kata dalam bentuk vektor

yang berdimensi rendah untuk menangkap hubungan semantik antar kata [13].

G. *Convolutional Neural Network* (CNN)

Convolutional Neural Network (CNN) merupakan salah satu arsitektur *deep learning* yang banyak digunakan untuk pemrosesan data dalam bentuk gambar atau teks. CNN awalnya dikembangkan untuk pengenalan citra dan pemrosesan gambar karena kemampuannya dalam mendeteksi fitur-fitur spasial melalui lapisan konvolusi, namun kini CNN juga diterapkan pada data teks, termasuk sentimen analisis berbasis aspek [14].

H. *Confusion matrix*

Confusion matrix pada dasarnya berisi data yang membandingkan hasil klasifikasi aktual dengan hasil yang diharapkan oleh sistem. Kinerja dari sistem klasifikasi jenis ini umumnya diukur melalui data yang terdapat dalam matriks tersebut [15].

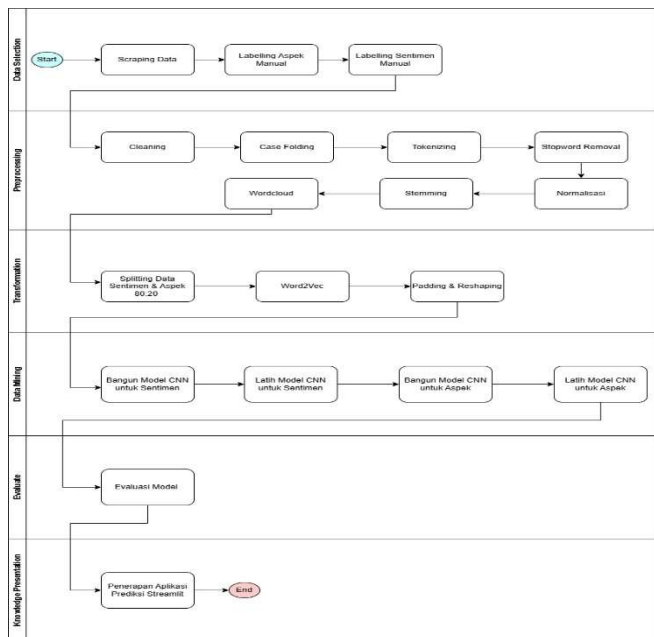
I. Streamlit

Streamlit merupakan sebuah *framework* berbasis *open-source* dari python yang dirancang untuk memudahkan dalam pengembangan aplikasi web secara interaktif, khususnya untuk keperluan *data science* dan juga *machine learning*. Dengan sintaksi yang sederhana, streamlit memungkinkan pengembang untuk membangun antaruka pengguna tanpa memerlukan keahlian mendalam terkait pengembangan web [16].

II. METODE

A. Tahapan Penelitian

Dalam penelitian ini, sistematika penyelesaian masalah adalah gambaran alur yang diperlukan dalam implementasi analisis sentimen berbasis aspek (*ABSA*) pada aplikasi Sekolah.mu. Penelitian ini menggunakan metode *Knowledge Discovery in Database* (KDD). Proses ini melibatkan beberapa tahap, seperti pemilihan data, pembersihan data, transformasi data, eksplorasi data (*data mining*), evaluasi hasil dan yang terakhir yaitu *knowledge presentation* melakukan prediksi melalui aplikasi streamlit.



GAMBAR 1
SISTEMATIKA PENYELESAIAN MASALAH

B. Pengumpulan Data

Data ulasan atau *review* dari pengguna aplikasi Sekolah.mu dikumpulkan melalui situs Google Play store dengan menggunakan teknik *scraping*. Data yang telah dikumpulkan dapat disimpan dalam format (CSV) untuk memudahkan analisis data lebih lanjut.

C. Pengolahan Data

Berikut merupakan langkah langkah yang akan dilakukan dalam tahapan pengolahan data:

1. *Data cleaning* adalah proses membersihkan data dari ketidaksesuaian, kesalahan, atau anomali untuk memastikan data siap digunakan dalam analisis atau visualisasi (Putri, 2024)
2. *Case Folding* kata adalah proses mengubah huruf kapital dalam sebuah data menjadi huruf kecil.
3. *Tokenization* adalah proses memecah teks menjadi unit-unit kecil yang disebut token, seperti kata, frasa, atau simbol (Parista, 2015)
4. *Stopword removal* adalah proses menghapus kata-kata yang dianggap tidak memiliki makna signifikan atau tidak relevan dalam analisis teks [17].
5. Normalisasi adalah tahapan untuk memperbaiki kata-kata yang salah penulisan, singkatan dan kata yang tidak baku.
6. *Stemming* adalah proses mengubah kata menjadi bentuk dasarnya dengan cara menghilangkan imbuhan seperti awalan, akhiran, atau sisipan [18].

D. Evaluasi

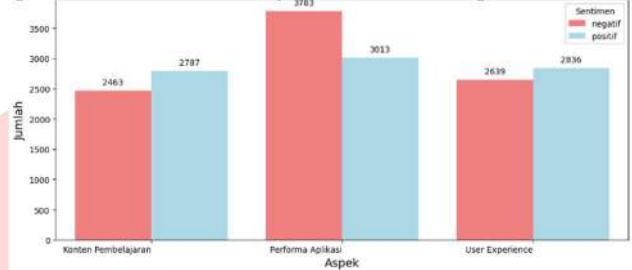
Pada penelitian ini, *confusion matrix* merupakan alat evaluasi yang akan digunakan. *Confusion matrix* merupakan metode evaluasi yang dapat menyatakan seberapa banyak

data yang dapat diuji, dan juga seberapa banyak variabel yang terdapat dalam tabel tersebut.

III. HASIL DAN PEMBAHASAN

A. Distribusi Data

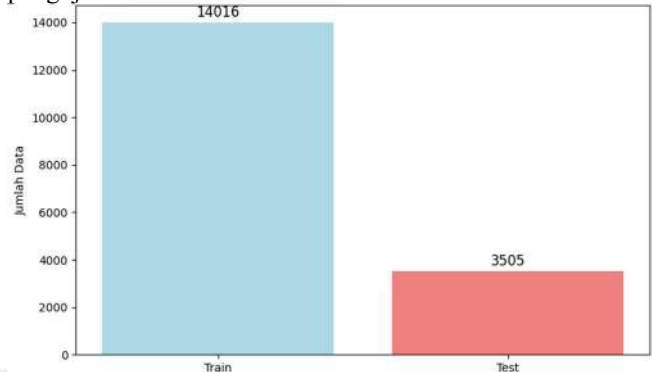
Dataset yang digunakan dalam penelitian ini berisi 17.521 baris data dengan tiga kolom atribut. Dataset ini merupakan hasil *scraping* yang telah diberikan label berdasarkan *egative* dan aspek. Label *egative* dikategorikan dengan dua label yaitu positif dan *egative*, sedangkan label aspek dikategorikan dalam tiga label yang mencakup Performa Aplikasi, Konten Pembelajaran, dan User Experience (UX).



GAMBAR 2
(DISTRIBUSI SENTIMEN POSITIF DAN NEGATIF TIAP ASPEK)

B. Pembagian Data Latih dan Data Uji

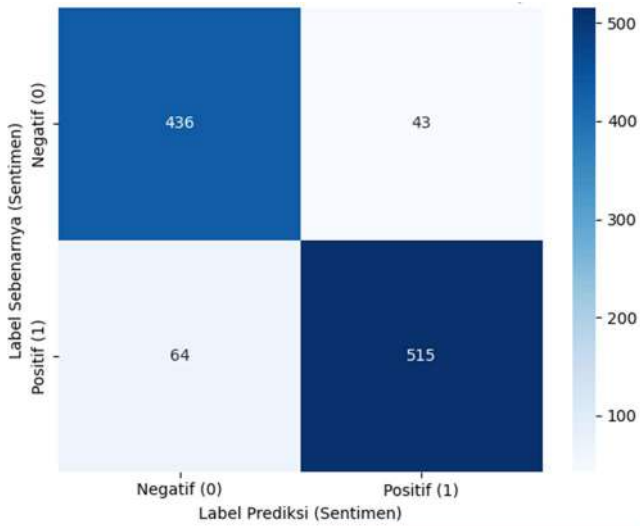
Dalam penelitian ini, pembagian data dilakukan dengan menggunakan proporsi 80% untuk pelatihan dan 20% untuk pengujian.



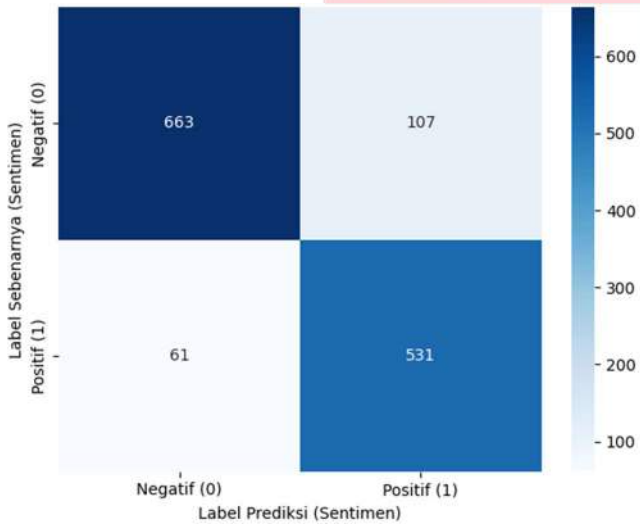
GAMBAR 3
(DISTRIBUSI DATA LATIH DAN DATA UJI)

C. Evaluation

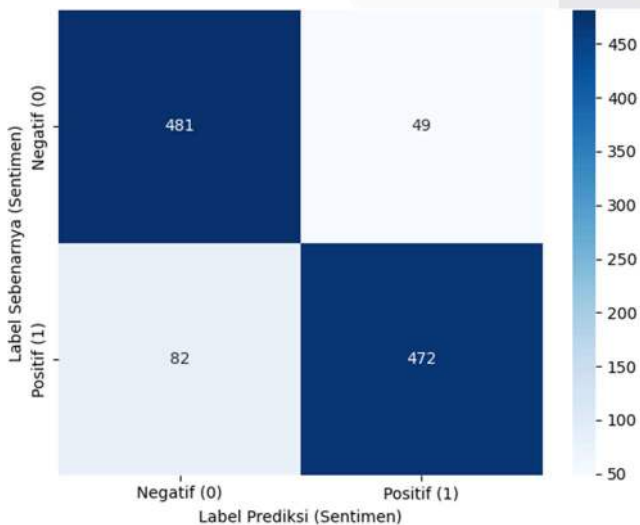
Confusion matrix memberikan gambaran mengenai performa model dalam pengklasifikasian data ke dalam tiga kelas aspek. Berikut merupakan hasil dari *confusion matrix* dan *classification report* yang dihasilkan dalam tiga aspek:



GAMBAR 4
(Confusion Matrx Aspek Konten Pembelajaran)



GAMBAR 5
(Confusion Matrix Aspek Performa Aplikasi)



GAMBAR 6
(Confusion Matrix Aspek User Experience)

Berikut merupakan tabel *classification report* untuk ketiga aspek:

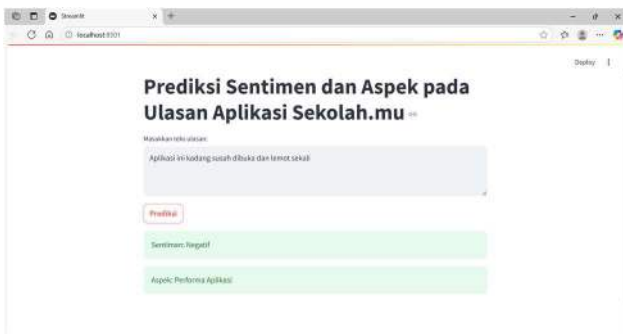
TABEL 1.
Classification Report Pada Ketiga Aspek

Classification Report Aspek Konten Pembelajaran				
Class	Precision	Recall	F1-Score	Support
0	0.87	0.91	0.89	479
1	0.92	0.89	0.91	579
Accuracy				1058
Macro avg	0.90	0.90	0.90	1058
Weight avg	0.90	0.90	0.90	1058
Classification Report Aspek Performa Aplikasi				
Class	Precision	Recall	F1-Score	Support
0	0.92	0.86	0.89	770
1	0.83	0.90	0.86	592
Accuracy				1362
Macro avg	0.88	0.88	0.88	1362
Weight avg	0.88	0.88	0.88	1362
Classification Report Aspek User Experience				
Class	Precision	Recall	F1-Score	Support
0	0.85	0.91	0.88	530
1	0.91	0.85	0.88	554
Accuracy				1084
Macro avg	0.88	0.88	0.88	1084
Weight avg	0.88	0.88	0.88	1084

Berdasarkan hasil *classification report*, dapat dianalisis bahwa *convolutional neural network* (CNN) menunjukkan performa yang baik dalam klasifikasi ulasan untuk tiga aspek, yaitu konten pembelajaran, performa aplikasi, dan *user experience*. Aspek performa aplikasi dan *user experience*, model CNN mengalami sedikit kesalahan klasifikasi terhadap kelas-kelas tertentu yang menyebabkan penurunan recall dan F1-score. Meskipun demikian, model tetap menunjukkan kinerja yang *solid* dengan hasil yang sesuai dalam identifikasi ulasan terhadap ketiga aspek. Maka dari itu, model CNN dapat digunakan sebagai model klasifikasi yang efektif untuk mengklasifikasikan ulasan dalam tiga aspek konten pembelajaran, performa aplikasi, dan *user experience*.

D. Penerapan Streamlit untuk Prediksi

Dalam mengimplementasikan model klasifikasi sentimen berbasis aspek, penulis membangun aplikasi prediksi berbasis web menggunakan *framework* streamlit. Aplikasi ini memiliki antarmuka yang sederhana yang memungkinkan pengguna untuk memasukkan teks ulasan secara langsung ke dalam kolom input. Setelah tombol prediksi diklik, maka sistem akan secara langsung menampilkan hasil prediksi yang terdiri dari sentimen positif dan negatif dan juga aspek seperti konten pembelajaran, performa aplikasi atau *user experience* yang terdeteksi dari ulasan tersebut.



GAMBAR 7
(Tampilan Aplikasi Prediksi Streamlit)

E. Analisis Sentimen Berbasis Aspek

Setiap aspek dilakukan analisis sentimen untuk memahami respon dari pengguna di berbagai fitur dalam aplikasi Sekolah.mu. Sebagai visualisasi, penelitian ini menggunakan *word cloud* untuk memberikan gambaran kata yang paling sering muncul dalam ulasan, memberikan *insight* mengenai topik atau fitur yang sering dibahas oleh pengguna terkait aspek tersebut.



GAMBAR 8
(Word Cloud pada Aspek Konten Pembelajaran)

Melalui visualisasi pada aspek performa aplikasi, secara umum dinilai mudah dan menyenangkan oleh banyak pengguna. Akan tetapi, untuk aspek pendukung layanan yang berkaitan dengan konten seperti pengelolaan sertifikat dan respon terhadap masalah teknis masih menjadi titik lemah yang perlu diperbaiki untuk mengoptimalkan pengalaman belajar oleh pengguna.



GAMBAR 9
(Word Cloud pada Aspek Performa Aplikasi)

Melalui visualisasi ini, dapat terlihat bahwa meskipun banyak pengguna merasa terbantu dengan kinerja aplikasi dalam menjalankan fungsinya, masih ditemukan keluhan teknis yang cukup signifikan, terutama terkait akses kelas dan kecepatan respons. Hal ini menunjukkan perlunya optimalisasi performa, khususnya untuk proses penting seperti loading materi, instalasi, dan navigasi antarmuka.



Gambar 10
(Word cloud pada Aspek User Experience)

Melalui visualisasi pada aspek user experience, terlihat bahwa meskipun fitur dan performa aplikasi secara umum dipapresiasi oleh pengguna, namun masih terdapat hambatan terkait elemen desain antarmuka dan kemudahan penggunaan. Beberapa pengguna mengalami kebingungan dalam penggunaan tombol dan fitur tertentu yang dinilai kurang intuitif. Perbaikan pada aspek desain dan navigasi dapat meningkatkan pengalaman pengguna secara menyeluruh.

F. Rekomendasi Terhadap Pihak Sekolah.mu

Berdasarkan hasil dari analisis sentimen berbasis aspek pada ulasan pengguna aplikasi Sekolah.mu, berikut beberapa rekomendasi yang dapat diberikan kepada pihak Sekolah.mu dalam peningkatan kualitas aplikasi dan pengalaman pengguna:

1. Hasil penelitian menunjukkan bahwa ulasan negatif khususnya pada konten pembelajaran lebih berfokus kepada keluhan terkait pelayanan pelanggan dan permasalahan dalam mengakses sertifikat. Untuk itu, Sekolah.mu disarankan untuk memperbaiki *service level agreement (SLA)* pada respons layanan pengguna, mempercepat waktu respon terhadap permasalahan teknis, serta mempermudah proses pengunduhan sertifikat bagi para pemakai sistem.
2. Meskipun banyak pengguna merasa puas dengan stabilitas aplikasi, ada keluhan seputar aksesibilitas kelas serta masalah yang terkait dengan pemasangan aplikasi. Oleh karena itu, Sekolah.mu perlu fokus pada pengoptimalan untuk proses sistem aplikasi seperti pemasangan dan pemuatan materi pembelajaran. Pengujian lebih lanjut terhadap kinerja aplikasi di berbagai platform dan perangkat dapat membantu mengidentifikasi kesenjangan terkait yang memerlukan pengoptimalan dalam kecepatan respons, mengatasi malfungsi teknis sekaligus meningkatkan responsivitas dan mengurangi masalah.
3. Analisis sentimen terhadap user experience menunjukkan keluhan berkaitan dengan navigasi serta desain antarmuka aplikasi yang Dirasa tidak intuitif. Beberapa pengguna bahkan melaporkan ada tombol dan fitur yang sulit untuk digunakan, tanpa syarat pengguna bisa mendapatkan pengalaman pengguna yang maksimal. Untuk itu, disarankan dilakukan perbaikan pada desain antarmuka aplikasinya agar lebih bersahabat kepada pengguna, berorientasi kepada kemudahan navigasi dan penempatan tombol yang lebih intuitif. Melakukan pengujian terhadap desain baru dengan melibatkan pengguna dapat menjadi langkah penting dalam meningkatkan pengalaman penggunaan aplikasi.

E. KESIMPULAN

Berdasarkan penelitian ini, akurasi klasifikasi sentimen positif dan negatif pada tiga aspek utama aplikasi seperti konten pembelajaran, performa aplikasi, serta *user experience* serta diperoleh antara 88% hingga 90% dengan menggunakan

algoritma *Convolutional Neural Network* (CNN). Hasil analisis menunjukkan pada konten pembelajaran evaluasi positif lebih berfokus kepada kemudahan materi yang disampaikan, sedangkan evaluasi negatif lebih banyak mengeluhkan *customer service* dan akses terhadap sertifikat. Pada performa aplikasi pendapat positif mencerminkan kepuasan terhadap kestabilan aplikasi, sementara kritik sebagian besar ditujukan kepada isu-isu teknis. Evaluasi *user experience* menunjukkan pengguna dengan suara positif menghargai fitur dan *interface* aplikasi untuk antarmuka program namun pengulas negatif turut memberikan komentar tentang navigasi dan desain yang tidak intuitif. Penyajian model *Convolutional Neural Network* (CNN) untuk klasifikasi sentimen dan aspek berbasis ulasan pengguna aplikasi Sekolah.mu mampu diterapkan dengan menggunakan aplikasi prediksi berbasis *streamlit*. Aplikasi ini memungkinkan pengguna memasukkan ulasan dengan mudah dan secara langsung menerima hasil prediksi.

REFERENSI

- [1] A. Simanungkalit, J. P. P. Naibaho, and A. De Kweldju, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Shopee Menggunakan Algoritma Naïve Bayes," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, p. 659, 2024, doi: 10.35889/jutisi.v13i1.1826.
- [2] H. Faisal, A. Febriandirza, and F. N. Hasan, "Analisis Sentimen Terkait Ulasan Pada Aplikasi PLN Mobile Menggunakan Metode Support Vector Machine," *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 5, no. 1, pp. 303–312, 2024.
- [3] I. M. Karo Karo, S. Dewi, and A. Perdana, "Implementasi Text Summarization Pada Review Aplikasi Digital Library System Menggunakan Metode Maximum Marginal Relevance," *JEKIN - J. Tek. Inform.*, vol. 4, no. 1, pp. 25–31, 2024, doi: 10.58794/jekin.v4i1.671.
- [4] H. Chyntia Morama, D. E. Ratnawati, and I. Arwani, "Analisis Sentimen berbasis Aspek terhadap Ulasan Hotel Tentrem Yogyakarta menggunakan Algoritma Random Forest Classifier," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 4, pp. 1702–1708, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [5] H. Mustakim and S. Priyanta, "Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 16, no. 2, p. 113, 2022, doi: 10.22146/ijccs.68903.
- [6] Nurul Hidayati, Faqih Hamami, and Riska Yanu Fa'rifah, "Aspect-Based Sentiment Analysis On FLIP Application Reviews (Play Store) Using Support Vector Machine (SVM) Algorithm," *J. Informatics Telecommun. Eng.*, vol. 7, no. 1, pp. 183–197, 2023, doi: 10.31289/jite.v7i1.9768.
- [7] P. A. Aritonang, M. E. Johan, and I. Prasetiawan, "Aspect-Based Sentiment Analysis on Application Review using CNN (Case Study : Peduli Lindungi Application)," *Ultim. Infosys J. Ilmu Sist. Inf.*, vol. 13, no. 1, pp. 54–61, 2022.
- [8] N. Nurwanda, N. Suarna, and W. Prihartono, "Penerapan Nlp (Natural Language Processing) Dalam Analisis Sentimen Pengguna Telegram Di Playstore," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1841–1846, 2024, doi: 10.36040/jati.v8i2.8469.
- [9] N. Nurzaman, N. Suarna, and W. Prihartono, "Analisis Sentimen Ulasan Aplikasi Threads Di Google Playstore Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 967–974, 2024, doi: 10.36040/jati.v8i1.8708.
- [10] C. A. Bahri and L. H. Suadaa, "Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 17, no. 1, p. 79, 2023, doi: 10.22146/ijccs.77354.
- [11] A. Agung, A. Daniswara, I. Kadek, and D. Nuryana, "Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru," *J. Informatics Comput. Sci.*, vol. 05, pp. 97–100, 2023.
- [12] Y. A. Prasetyo, E. Utami, and A. Yaqin, "Pengaruh Komposisi Split Data Terhadap Performa Akurasi Analisis Sentimen Algoritma Naïve Bayes dan SVM," vol. 6, no. 2, pp. 382–390, 2024, doi: 10.33650/jeeecom.v4i2.
- [13] D. Iskandar and A. Kurniawati, "A COMPARATIVE ANALYSIS OF WORD2VEC AND DOC2VEC TECHNIQUES IN," vol. 12, no. 1, pp. 133–144, 2025, doi: 10.25126/jtiik.2025129143.
- [14] M. T. Ari Bangsa, S. Priyanta, and Y. Suyanto, "Aspect-Based Sentiment Analysis of Online Marketplace Reviews Using Convolutional Neural Network," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 14, no. 2, p. 123, 2020, doi: 10.22146/ijccs.51646.
- [15] a. K. Santra and C. J. Christy, "Genetic Algorithm and Confusion Matrix for Document Clustering," *Int. J. Comput. Sci.*, vol. 9, no. 1, pp. 322–328, 2012, [Online]. Available: <http://ijesi.org/papers/IJCSI-9-1-2-322-328.pdf>
- [16] M. Ferdandi, N. Y. Setiawan, and F. Abdurrachman Bachtiar, "Prediksi Potensi Penjualan Makanan Beku Berdasarkan Ulasan Pengguna Shopee Menggunakan Metode Decision Tree Algoritma C4.5 Dan Random Forest (Studi Kasus Dapur Lilis)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 2, pp. 588–596, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [17] R. Rinandyaswara, Y. A. Sari, and M. T. Furqon, "Pembentukan Daftar Stopword Menggunakan Term Based Random Sampling Pada Analisis Sentimen Dengan Metode Naïve Bayes (Studi Kasus: Kuliah Daring Di Masa Pandemi)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 4, p. 717, 2022, doi: 10.25126/jtiik.2022934707.
- [18] M. U. Albab, Y. Karuniawati P, and M. N. Fawaiq,

“Optimization of the Stemming Technique on Text preprocessing President 3 Periods Topic,” *J. Transform.*, vol. 20, no. 2, pp. 1–10, 2023, [Online]. Available: <https://journals.usm.ac.id/index.php/transformatika/> ■page1

