

Aspect-Based Sentiment Analysis Penggunaan ChatGPT dalam Pendidikan: Perbandingan Model LSTM, Bi-LSTM, dan CNN

Dita Yulianti¹⁾, Novera Kristianti²⁾, Agus Sehatman Saragih³⁾, Tomas Leonardo⁴⁾

¹⁾²⁾³⁾⁴⁾Jurusan Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya
Kampus Tanjung Nyaho, Jl. Yos Sudarso, Palangka Raya

¹⁾ydita97@gmail.com

²⁾noverakristianti@eng.upr.ac.id

³⁾assaragih@it.upr.ac.id

⁴⁾tomasleonardo@it.upr.ac.id

Abstrak

Perkembangan teknologi kecerdasan buatan telah membawa perubahan besar dalam berbagai bidang, termasuk pendidikan. Salah satu teknologi yang banyak digunakan adalah ChatGPT, sebuah model bahasa berbasis AI yang dimanfaatkan oleh pelajar dan mahasiswa untuk mendukung aktivitas akademik. Meskipun pemanfaatannya semakin luas, opini publik terhadap penggunaannya dalam dunia pendidikan masih beragam. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap ChatGPT berdasarkan aspek tertentu, serta mengevaluasi kinerja beberapa algoritma *deep learning* dalam klasifikasi teks.

Penelitian ini menggunakan pendekatan *Aspect-Based Sentiment Analysis* (ABSA), dengan pemodelan aspek melalui metode *Latent Dirichlet Allocation* (LDA) dan pelabelan sentimen menggunakan IndoBERT. Data dikumpulkan dari platform X (Twitter) dengan teknik *scraping*, lalu diproses melalui tahapan *preprocessing*, ekstraksi fitur, pembagian data, penerapan SMOTE, dan evaluasi model. Tiga algoritma klasifikasi yang digunakan adalah *Long Short-Term Memory* (LSTM), *Bidirectional LSTM* (Bi-LSTM), dan *Convolutional Neural Network* (CNN).

Hasil analisis menunjukkan bahwa opini publik terhadap ChatGPT dalam pendidikan cenderung positif, terutama pada aspek *Asisten Belajar* (62,1%) dan *Etika Akademik* (55,3%). Pada klasifikasi aspek, model Bi-LSTM dengan *split* data 90:10 mencatat akurasi tertinggi sebesar 91,11%. Sementara itu, pada klasifikasi sentimen, model yang sama memperoleh akurasi sebesar 87,88% pada *split* data yang sama. Temuan ini menunjukkan bahwa Bi-LSTM memiliki performa yang baik dalam kedua jenis tugas klasifikasi. Penelitian ini memberikan gambaran mengenai persepsi publik terhadap penggunaan ChatGPT di bidang pendidikan serta mendukung pemilihan model yang sesuai untuk analisis sentimen berbasis aspek.

Kata kunci: ChatGPT, Sentimen, *Aspect-Based Sentiment Analysis*, Pendidikan

Abstract

The development of artificial intelligence technology has brought significant changes across various fields, including education. One widely used technology is ChatGPT, an AI-based language model utilized by students to support academic activities. Despite its growing adoption, public opinion on its use in education remains diverse. This study aims to analyze public sentiment toward ChatGPT based on specific aspects and evaluate the performance of several deep learning algorithms in text classification.

This research employs an Aspect-Based Sentiment Analysis (ABSA) approach, with aspect modeling conducted using the Latent Dirichlet Allocation (LDA) method and sentiment labeling performed using IndoBERT. Data were collected from the X platform (formerly Twitter) using scraping techniques and processed through stages preprocessing, feature extraction, data splitting, SMOTE application, and model evaluation. Three classification algorithms were used: Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and Convolutional Neural Network (CNN).

The analysis results show that public opinion toward ChatGPT in education tends to be positive, particularly in the aspects of Learning Assistance (62.1%) and Academic Ethics (55.3%). In aspect classification, the Bi-LSTM model with a 90:10 data split recorded the highest accuracy of 91.11%. Meanwhile, in sentiment classification, the same model achieved an accuracy of 87.88% with the same data split. These findings indicate that Bi-LSTM performs well in both types of classification tasks. This study provides insight into public perceptions of ChatGPT usage in education and supports the selection of appropriate models for aspect-based sentiment analysis.

Keywords: ChatGPT, Sentiment, Aspect-Based Sentiment Analysis, Education

1. PENDAHULUAN

Perkembangan teknologi kecerdasan buatan (AI) telah memberikan dampak signifikan dalam dunia pendidikan, khususnya melalui pemanfaatan model bahasa seperti ChatGPT. Sejak diperkenalkan pada November 2022, ChatGPT semakin banyak digunakan oleh pelajar dan mahasiswa untuk menyusun esai, memahami konsep akademik, dan mencari informasi secara interaktif [1]. Namun, adopsi teknologi ini juga menimbulkan perdebatan terkait keandalan informasi serta pengaruhnya terhadap keterampilan berpikir kritis dan menulis. Survei Pew Research Center menunjukkan peningkatan penggunaan ChatGPT di kalangan pelajar Amerika Serikat dari 13% (2023) menjadi 26% (2024), meskipun hanya 18% yang menganggapnya layak digunakan untuk menulis esai [2]. Di Indonesia, tren serupa mulai terlihat, namun persepsi publik masih beragam. Penelitian Pratama & Hendry [3] menunjukkan bahwa sentimen terhadap ChatGPT di platform X (Twitter) cenderung netral, sementara studi Ramaputra & Purnomo [4] lebih menekankan pada evaluasi algoritma analisis sentimen.

Untuk memahami opini publik secara lebih mendalam, penelitian ini menerapkan pendekatan Aspect-Based Sentiment Analysis (ABSA), yang memungkinkan identifikasi sentimen berdasarkan kategori aspek tertentu [5]. Platform X dipilih sebagai sumber data karena digunakan secara luas di Indonesia untuk menyampaikan opini dan aspirasi [6]. Metode yang digunakan melibatkan ekstraksi fitur teks menggunakan IndoBERT [7], yang kemudian diklasifikasikan menggunakan tiga model: Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), dan Convolutional Neural Network (CNN). LSTM digunakan untuk menangkap dependensi jangka panjang dalam teks [8], Bi-LSTM sebagai pengembangan dari LSTM [9], dan CNN untuk mengenali pola lokal dalam data teks [10]. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap penggunaan ChatGPT dalam pendidikan berdasarkan aspek-aspek tertentu, serta membandingkan kinerja ketiga model klasifikasi tersebut dalam tugas ABSA.

2. TINJAUAN PUSTAKA

2.1 Aspect-Based Sentiment Analysis

Aspect-Based Sentiment Analysis (ABSA) dikembangkan untuk mengatasi keterbatasan analisis sentimen tradisional yang tidak membedakan opini terhadap aspek spesifik dalam teks. ABSA memungkinkan identifikasi aspek yang dibahas serta sentimen yang menyertainya dalam satu ulasan [5].

2.2 LSTM

Long Short-Term Memory (LSTM) merupakan arsitektur jaringan saraf yang efektif dalam menangani data sekuensial, seperti analisis sentimen, karena kemampuannya menangkap ketergantungan jangka panjang dalam teks [7]. LSTM terdiri dari memory cell dan tiga gerbang utama yaitu *input*, *forget*, dan *output gate* yang mengatur aliran informasi secara dinamis [9].

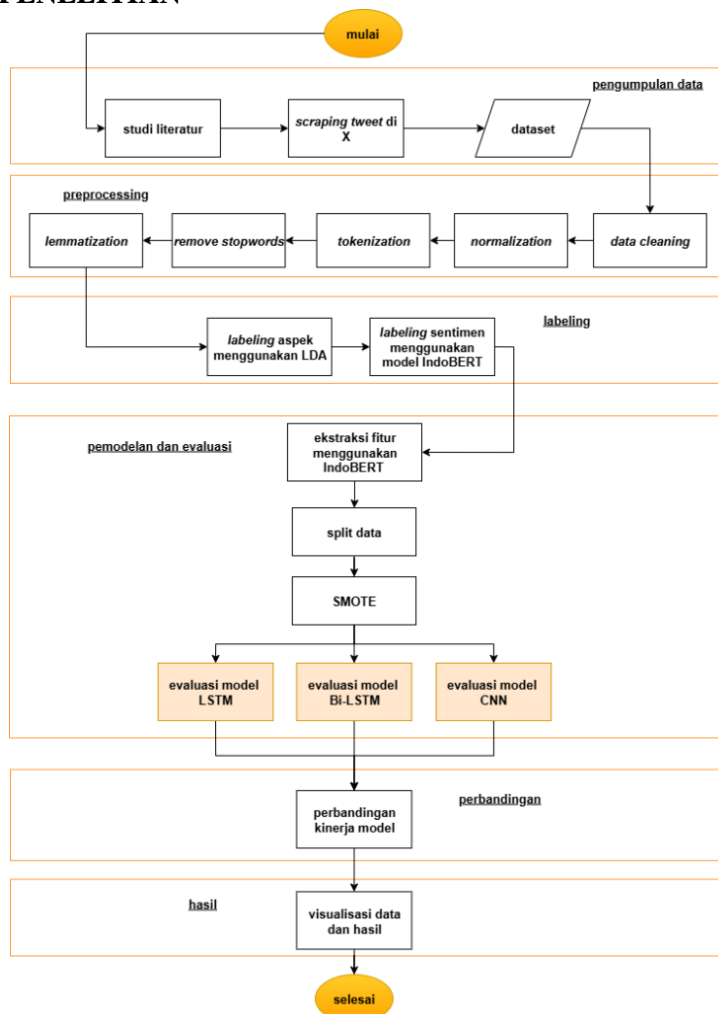
2.3 Bi-LSTM

Bidirectional Long Short-Term Memory (Bi-LSTM) merupakan pengembangan dari LSTM yang memproses data sekuensial dalam dua arah maju dan mundur sehingga mampu menangkap konteks sebelum dan sesudah suatu kata dalam kalimat [11]. Dengan mekanisme ini, Bi-LSTM menghasilkan representasi yang lebih menyeluruh dan meningkatkan akurasi dalam tugas klasifikasi.

2.4 CNN

Convolutional Neural Network (CNN) awalnya dikembangkan untuk pemrosesan citra, namun juga efektif dalam klasifikasi teks seperti analisis sentimen. Dalam konteks ini, teks diubah menjadi representasi numerik misalnya melalui BERT, sebelum diproses oleh CNN sebagai pengklasifikasi [7]. Arsitektur CNN untuk teks umumnya terdiri dari embedding layer, lapisan konvolusi satu dimensi (Conv1D), *global max pooling*, *dense layer*, dan *output layer*, yang bekerja berurutan untuk mengekstraksi fitur dan menghasilkan klasifikasi akhir [12].

3. METODE PENELITIAN



Gambar 1. Alur penelitian

Penelitian ini diawali dengan proses *scraping* data dari *platform X*, dilanjutkan tahap *preprocessing* berupa pembersihan data, normalisasi, tokenisasi, penghapusan *stopwords*, dan *lemmatization*. Aspek ditentukan menggunakan LDA, sedangkan pelabelan sentimen dilakukan dengan IndoBERT. Fitur teks diekstraksi menggunakan IndoBERT, lalu data dibagi dalam rasio 70:30, 80:20, dan 90:10. Pemodelan dilakukan menggunakan LSTM, Bi-LSTM, dan CNN, kemudian dievaluasi menggunakan akurasi, presisi, *recall*, dan *F1-score*. Hasil akhir dibandingkan dan divisualisasikan.

4. PEMBAHASAN

4.1 Pengumpulan Data

Pengumpulan data dilakukan melalui metode *web scraping* terhadap platform media sosial X menggunakan *tool Tweet Harvest* pada *tweet* berbahasa Indonesia terkait penggunaan ChatGPT dalam pendidikan, dengan rentang waktu Januari 2023-April 2025 dan kata kunci seperti “chatgpt belajar” dan “chatgpt akademik”. Diperoleh data sebanyak 14.138 *tweet*.

Tabel 1. Dataset

<i>created_at</i>	<i>full_text</i>
Tue Jun 13 13:58:06 +0000 2023	terima kasih kepada chatGPT ak bisa belajar dengan mudah dan nyaman
Thu Mar 27 03:06:48 +0000 2025	Alasan kampus melarang menggunakan AI seperti ChatGPT dkk. 1. Plagiarisme dan keaslian karya 2. Kurangnya pemahaman mendalam 3. Akurasi dan validitas data 4. Etika akademik 5. Kurangnya keterampilan riset dan analisis #fyp

4.2 Preprocessing

4.2.1 Cleaning

Data cleaning dilakukan dengan menghapus kolom tidak relevan, *tweet* duplikat, elemen non-teks seperti *URL*, *mention*, dan *emoji*, serta mengubah teks menjadi *lowercase*. Hasil akhir berjumlah 8.660 *tweet* bersih.

Tabel 2. Hasil *data cleaning*

<i>full_text</i>	<i>casefolded_fulltext</i>
terima kasih kepada chatGPT ak bisa belajar dengan mudah dan nyaman	terima kasih kepada chatgpt ak bisa belajar dengan mudah dan nyaman
Alasan kampus melarang menggunakan AI seperti ChatGPT dkk. 1. Plagiarisme dan keaslian karya 2. Kurangnya pemahaman mendalam 3. Akurasi dan validitas data 4. Etika akademik 5. Kurangnya keterampilan riset dan analisis #fyp	alasan kampus melarang menggunakan ai seperti chatgpt dkk plagiarisme dan keaslian karya kurangnya pemahaman mendalam akurasi dan validitas data etika akademik kurangnya keterampilan riset dan analisis fyp

4.2.2 Normalization

Normalization dilakukan dengan mengganti kata tidak baku menggunakan kamus *slang* digital, sehingga teks menjadi lebih konsisten dan formal.

Tabel 3. Hasil normalisasi

<i>casefolded_fulltext</i>	<i>normalized_fulltext</i>
terima kasih kepada chatgpt ak bisa belajar dengan mudah dan nyaman	terima kasih kepada chatgpt saya bisa belajar dengan mudah dan nyaman
alasan kampus melarang menggunakan ai seperti chatgpt dkk plagiarisme dan keaslian karya kurangnya pemahaman mendalam akurasi dan validitas data etika akademik kurangnya keterampilan riset dan analisis fyp	alasan kampus melarang menggunakan artificial intelligence seperti chatgpt plagiarisme dan keaslian karya kurangnya pemahaman mendalam akurasi dan validitas data etika akademik kurangnya keterampilan riset dan analisis

4.2.3 Tokenization

Tokenisasi bertujuan untuk memecah teks menjadi satuan kata atau *token* agar struktur bahasanya dapat dianalisis lebih mudah oleh sistem.

Tabel 4. Hasil tokenisasi

<i>normalized_fulltext</i>	<i>tokens</i>
terima kasih kepada chatgpt saya bisa belajar dengan mudah dan nyaman	['terima', 'kasih', 'kepada', 'chatgpt', 'saya', 'bisa', 'belajar', 'dengan', 'mudah', 'dan', 'nyaman']
alasan kampus melarang menggunakan artificial intelligence seperti chatgpt plagiarisme dan keaslian karya kurangnya pemahaman mendalam akurasi dan validitas data etika akademik kurangnya keterampilan riset dan analisis	['alasan', 'kampus', 'melarang', 'menggunakan', 'artificial', 'intelligence', 'seperti', 'chatgpt', 'plagiarisme', 'dan', 'keaslian', 'karya', 'kurangnya', 'pemahaman', 'mendalam', 'akurasi', 'dan', 'validitas', 'data', 'etika', 'akademik', 'kurangnya', 'keterampilan', 'riset', 'dan', 'analisis']

4.2.4 Remove Stopwords

Stopwords dihapus menggunakan pustaka *NLTK* untuk menyaring kata umum yang tidak relevan, kecuali kata negasi yang penting bagi analisis sentimen.

Tabel 5. Hasil remove stopwords

tokens	remove_stopwords
['terima', 'kasih', 'kepada', 'chatgpt', 'saya', 'bisa', 'belajar', 'dengan', 'mudah', 'dan', 'nyaman']	['terima', 'kasih', 'chatgpt', 'belajar', 'mudah', 'nyaman']
['alasan', 'kampus', 'melarang', 'menggunakan', 'artificial', 'intelligence', 'seperti', 'chatgpt', 'plagiarisme', 'dan', 'keaslian', 'karya', 'kurangnya', 'pemahaman', 'mendalam', 'akurasi', 'dan', 'validitas', 'data', 'etika', 'akademik', 'kurangnya', 'keterampilan', 'riset', 'dan', 'analisis']	['alasan', 'kampus', 'melarang', 'artificial', 'intelligence', 'chatgpt', 'plagiarisme', 'keaslian', 'karya', 'kurangnya', 'pemahaman', 'mendalam', 'akurasi', 'validitas', 'data', 'etika', 'akademik', 'kurangnya', 'keterampilan', 'riset', 'analisis']

4.2.5 Lemmatization

Lemmatization dilakukan untuk mengubah kata menjadi bentuk dasar sesuai kaidah bahasa Indonesia.

Tabel 6. Hasil lemmatisasi

remove_stopwords	lemmatized_tokens
['terima', 'kasih', 'chatgpt', 'belajar', 'mudah', 'nyaman']	['terima', 'kasih', 'chatgpt', 'belajar', 'mudah', 'nyaman']
['alasan', 'kampus', 'melarang', 'artificial', 'intelligence', 'chatgpt', 'plagiarisme', 'keaslian', 'karya', 'kurangnya', 'pemahaman', 'mendalam', 'akurasi', 'validitas', 'data', 'etika', 'akademik', 'kurangnya', 'keterampilan', 'riset', 'analisis']	['alasan', 'kampus', 'melarang', 'artificial', 'intelligence', 'chatgpt', 'plagiarisme', 'asli', 'karya', 'kurang', 'paham', 'dalam', 'akurasi', 'validitas', 'data', 'etika', 'akademik', 'kurang', 'terampil', 'riset', 'analisis']

4.3 Pelabelan Aspek

4.3.1 Persiapan Corpus

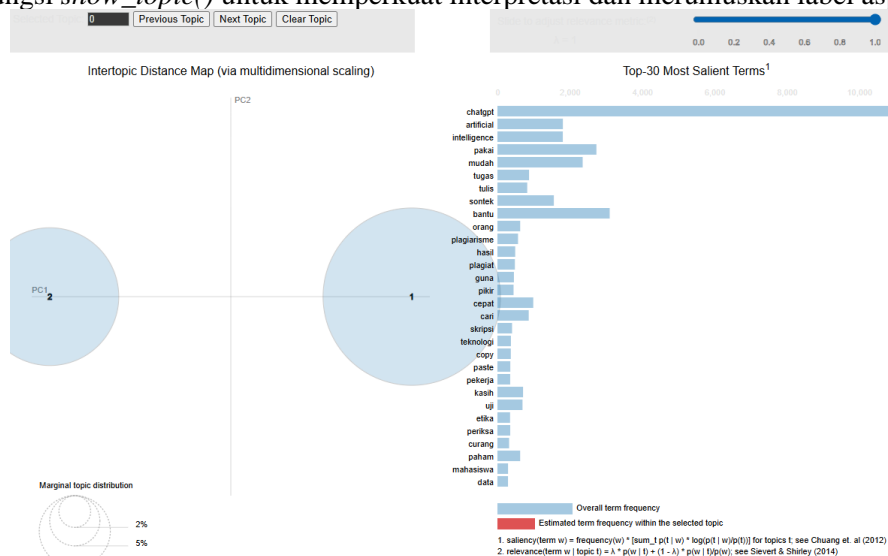
Corpus diubah jadi tokens, lalu dibuat dictionary dan Bag-of-Words sebagai input numerik untuk LDA.

4.3.2 Penentuan Jumlah Topik

Setelah membentuk Bag-of-Words, jumlah topik optimal ditentukan melalui kombinasi nilai koherensi, interpretasi manual, dan visualisasi *intertopic distance map* [13]. Awalnya diuji 10 topik, namun topik 3–8 menunjukkan tumpang tindih semantik. Analisis kata kunci mengindikasikan bahwa topik-topik tersebut membahas pengalaman belajar siswa. Eksplorasi dilanjutkan dengan 1–3 topik, dan model dua topik dipilih karena memiliki koherensi tertinggi (0,315) serta pemisahan visual yang jelas.

4.3.3 Visualisasi Topik

Visualisasi semantik topik dilakukan menggunakan *PyLDAvis*, yang menampilkan sebaran topik dalam bentuk *bubble* dan daftar kata kunci utama. Ekstraksi kata kunci tambahan dilakukan dengan fungsi *show_topic()* untuk memperkuat interpretasi dan merumuskan label aspek.



Gambar 2. Visualisasi PyLDAvis

4.3.4 Penentuan Topik Dominan

Setelah model LDA optimal diperoleh, topik dominan untuk tiap *tweet* ditentukan berdasarkan distribusi probabilitas, dengan topik bernilai tertinggi sebagai representasi utama.

4.3.5 Perumusan Label dengan Gen AI

Setelah kata kunci tiap topik diekstraksi, label aspek dirumuskan menggunakan *Generative AI* Gemini Flash. Model menerima *prompt* berisi kata kunci dan menghasilkan dua label: *Asisten Belajar* untuk topik penggunaan ChatGPT sebagai alat bantu, dan *Etika Akademik* untuk topik seputar plagiarisme dan integritas. Hasil ini sejalan dengan studi sebelumnya yang menyoroti manfaat dan tantangan penggunaan ChatGPT dalam pendidikan.

4.3.6 Pemetaan Label

Setiap *tweet* diberi label aspek berdasarkan ID topik dominan, yang dicocokkan dengan hasil perumusan aspek. Label disimpan dalam kolom *aspect* pada *dataset*.

Tabel 7. Label aspek pada tweet

<i>tweet</i>	<i>aspect</i>
terima kasih chatgpt belajar mudah nyaman	asisten belajar
alasan kampus melarang artificial intelligence chatgpt plagiarisme asli karya kurang paham dalam akurasi validitas data etika akademik kurang terampil riset analisis	etika akademik

4.4 Pelabelan Sentimen

4.4.1 Model

Analisis sentimen dilakukan dengan model *IndoBERT* melalui pustaka *transformers*, menggunakan gabungan *tweet* dan label aspek dalam kolom *text_for_sentiment* untuk mempertahankan konteks opini.

4.4.2 Implementasi Analisis Sentimen

Setelah model dimuat, *pipeline* dijalankan pada kolom *text_for_sentiment* untuk menghasilkan label sentimen dan skor, yang disimpan dalam kolom *sentiment* dan *score*.

Tabel 8. Label sentimen pada tweet

<i>tweet</i>	<i>aspect</i>	<i>sentiment</i>	<i>score</i>
terima kasih chatgpt belajar mudah nyaman	asisten belajar	positif	0.9992
alasan kampus melarang artificial intelligence chatgpt plagiarisme asli karya kurang paham dalam akurasi validitas data etika akademik kurang terampil riset analisis	etika akademik	negatif	0.9987

4.5 Evaluasi Model

Fitur diekstraksi dengan *IndoBERT* untuk klasifikasi aspek dan sentimen, lalu data dibagi (70:30, 80:20, 90:10), diseimbangkan dengan *SMOTE*, dan dioptimasi menggunakan *Bayesian hyperparameter tuning* via *Optuna* pada LSTM, Bi-LSTM, dan CNN.

4.5.1 Model Aspek

Tabel 9. Hasil evaluasi LSTM aspek

<i>Split</i>	<i>Accuracy (%)</i>	<i>Precision (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
70:30	89,30	89,80	89,30	89,51
80:20	90,30	89,87	90,30	90,00
90:10	89,84	90,42	89,84	90,07

Model LSTM menunjukkan performa stabil di semua skenario, dengan *accuracy* tertinggi 90,30% (split 80:20) dan *F1-score* tertinggi 90,07% (split 90:10). Perbedaan antar skema kecil, menandakan ketahanan model terhadap variasi data latih.

Tabel 10. Hasil evaluasi Bi-LSTM aspek

<i>Split</i>	<i>Accuracy (%)</i>	<i>Precision (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
70:30	90,45	90,15	90,45	90,27
80:20	89,61	89,98	89,61	89,77
90:10	91,11	90,99	91,11	91,05

Model Bi-LSTM menunjukkan performa konsisten di semua skenario, dengan *F1-score* tertinggi 91,05% dan *accuracy* 91,11% pada split 90:10, menandakan ketahanan terhadap variasi data latih.

Tabel 11. Hasil evaluasi CNN aspek

<i>Split</i>	<i>Accuracy (%)</i>	<i>Precision (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
70:30	88,68	90,34	88,68	89,24
80:20	89,43	89,80	89,43	89,59
90:10	90,53	90,53	90,53	90,53

Model CNN menunjukkan performa konsisten di semua skenario, dengan *F1-score* tertinggi 90,53% pada split 90:10. Perbedaan antar skema kecil, menandakan kestabilan klasifikasi aspek meskipun proporsi data latih bervariasi.

4.5.2 Model Sentimen

Tabel 12. Hasil evaluasi LSTM sentimen

<i>Split</i>	<i>Accuracy (%)</i>	<i>Precision (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
70:30	85,76	85,69	85,76	85,66
80:20	85,97	86,32	85,97	86,06
90:10	87,07	87,57	87,07	87,17

Model LSTM menunjukkan performa konsisten dalam klasifikasi sentimen, dengan *F1-score* tertinggi 87,17% pada split 90:10, perbedaan antar skenario kecil, menandakan adaptabilitas terhadap variasi data latih.

Tabel 13. Hasil evaluasi Bi-LSTM sentimen

<i>Split</i>	<i>Accuracy (%)</i>	<i>Precision (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
70:30	84,33	84,98	84,33	84,47
80:20	86,03	86,67	86,03	86,15
90:10	87,88	87,95	87,88	87,90

Model Bi-LSTM menunjukkan tren positif dalam klasifikasi sentimen, dengan *F1-score* tertinggi 87,90% pada split 90:10, peningkatan proporsi data latih konsisten meningkatkan akurasi, serta keseimbangan prediksi antar kelas.

Tabel 14. Hasil evaluasi CNN sentimen

<i>Split</i>	<i>Accuracy (%)</i>	<i>Precision (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
70:30	85,45	85,59	85,45	85,50
80:20	85,97	85,97	85,97	85,97
90:10	86,84	87,47	86,84	86,95

Model CNN menunjukkan peningkatan bertahap dalam klasifikasi sentimen, dengan *F1-score* tertinggi 86,95% pada split 90:10, tren metrik yang konsisten mencerminkan stabilitas dan presisi saat proporsi data latih diperbesar.

4.6 Perbandingan Model

Tabel 15. Perbandingan model aspek

<i>Split</i>	<i>Model</i>	<i>Accuracy (%)</i>	<i>Precision (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
70:30	LSTM	89,30	89,80	89,30	89,51
	Bi-LSTM	90,45	90,15	90,45	90,27
	CNN	88,68	90,34	88,68	89,24
80:20	LSTM	90,30	89,87	90,30	90,00
	Bi-LSTM	89,61	89,98	89,61	89,77
	CNN	89,43	89,80	89,43	89,59
90:10	LSTM	89,84	90,42	89,84	90,07
	Bi-LSTM	91,11	90,99	91,11	91,05
	CNN	90,53	90,53	90,53	90,53

Bi-LSTM menunjukkan performa terbaik dalam klasifikasi aspek, terutama pada split 90:10 dengan *accuracy* 91,11% dan *F1-score* 91,05%, berkat pemrosesan dua arah yang mampu menangkap makna implisit dan konteks semantik tersebar. *CNN* mencatat *F1-score* tinggi di split 90:10 (90,53%) karena kekuatannya mengenali frasa eksplisit, namun tetap di bawah *Bi-LSTM*.

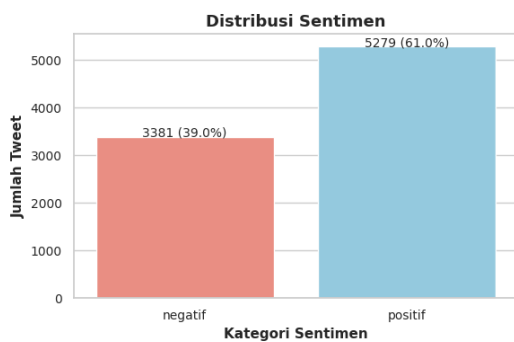
LSTM stabil di semua split, tetapi terbatas dalam memahami konteks non-linier akibat pemrosesan satu arah.

Tabel 16. Perbandingan model sentimen

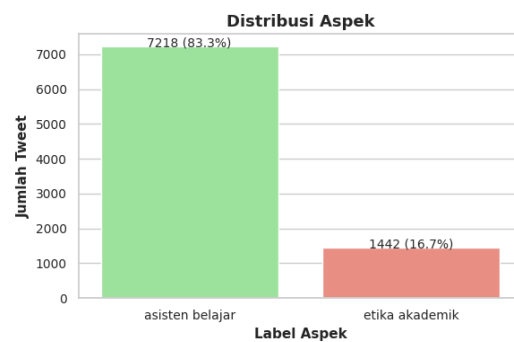
Split	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
70:30	LSTM	85,76	85,69	85,76	85,66
	Bi-LSTM	84,33	84,98	84,33	84,47
	CNN	85,45	85,59	85,45	85,50
80:20	LSTM	85,97	86,32	85,97	86,06
	BiLSTM	86,03	86,67	86,03	86,15
	CNN	85,97	85,97	85,97	85,97
90:10	LSTM	87,07	87,57	87,07	87,17
	BiLSTM	87,88	87,95	87,88	87,90
	CNN	86,84	87,47	86,84	86,95

Pada klasifikasi sentimen, *LSTM* mencatat performa terbaik di split 70:30 (*F1-score* 85,66%) berkat kemampuannya menangkap dependensi panjang dan sentimen implisit. Di split 80:20, *Bi-LSTM* unggul (*F1-score* 86,15%) karena pemrosesan dua arah yang efektif pada tweet kompleks. Pada split 90:10, *Bi-LSTM* kembali dominan dengan *accuracy* 87,88% dan *F1-score* 87,90%, menunjukkan kestabilan di distribusi data yang tidak seimbang. *CNN* mencatat skor cukup tinggi di semua skenario, namun cenderung fokus pada frasa eksplisit dan pola lokal, sehingga *precision* dan *recall* menurun saat menghadapi konteks semantik yang tersebar.

4.7 Visualisasi Data

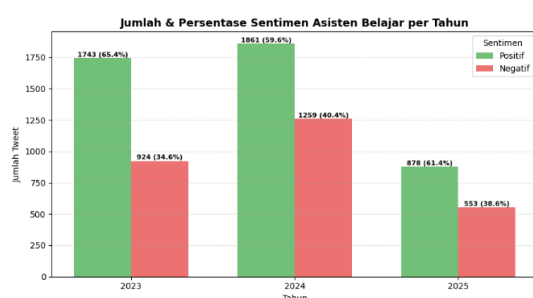


Gambar 3. Distribusi Sentimen

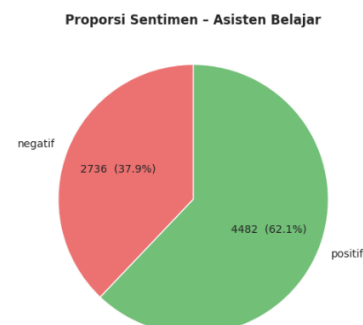


Gambar 4. Distribusi Aspek

Sentimen terhadap *ChatGPT* didominasi oleh tweet positif (61,0%), menunjukkan persepsi publik yang cenderung mendukung. Aspek *Asisten Belajar* paling banyak dibahas (83,3%), menandakan fokus utama pengguna adalah manfaat *ChatGPT* dalam pembelajaran, dibanding isu etika akademik.



Gambar 5. Tren sentimen aspek asisten belajar



Gambar 6. Proporsi sentimen aspek asisten belajar

Tren sentimen terhadap aspek *Asisten Belajar* menunjukkan dominasi opini positif dari tahun 2023 hingga 2025, dengan proporsi tertinggi di 2023 (65,4%) dan tetap stabil di 2025 (61,4%) meski volume tweet menurun. Secara keseluruhan, dari 7.218 tweet, sebanyak 4.482 (62,1%) bersentimen positif, memperkuat bahwa publik mengapresiasi *ChatGPT* sebagai alat bantu pembelajaran.

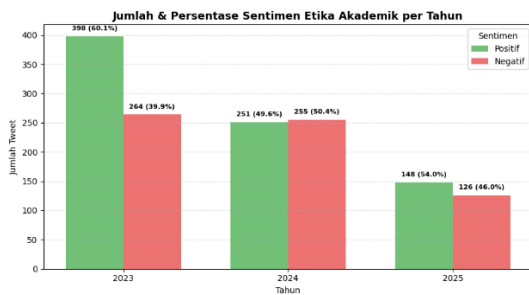


Gambar 7. Word cloud sentimen negatif aspek asisten belajar



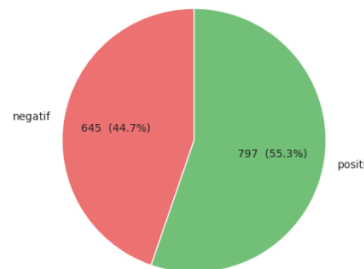
Gambar 8. Word cloud sentimen positif aspek asisten belajar

Word cloud sentimen negatif pada aspek *Asisten Belajar* menampilkan kata seperti “tidak,” “salah,” “malas,” dan “jangan,” mencerminkan kritik terhadap penggunaan *ChatGPT* dalam pembelajaran. Sebaliknya, word cloud positif menunjukkan kata “bantu,” “mudah,” “cepat,” dan “skripsi,” menandakan apresiasi pengguna terhadap kemudahan akses dan dukungan akademik yang diberikan oleh *ChatGPT*.



Gambar 9. Tren sentimen aspek etika akademik

Proporsi Sentimen - Etika Akademik



Gambar 10. Proporsi sentimen aspek asisten belajar

Sentimen terhadap aspek *Etika Akademik* menunjukkan tren relatif seimbang selama tiga tahun, dengan dominasi positif di 2023 (60,1%), penurunan di 2024 (49,6%), dan stabil di 2025 (54,0%). Dari total 1.442 tweet, sebanyak 797 (55,3%) bersentimen positif, menandakan apresiasi publik terhadap isu etika dalam penggunaan *ChatGPT* di lingkungan akademik.



Gambar 11. Word cloud sentimen negatif aspek etika akademik



Gambar 12. Word cloud sentimen positif aspek etika akademik

Word cloud negatif pada aspek *Etika Akademik* menampilkan kata seperti “plagiarisme,” “malas,” dan “tuliskan,” mencerminkan kekhawatiran terhadap kejujuran akademik dan penurunan usaha berpikir. Sebaliknya, word cloud positif menunjukkan kata “bantu,” “mudah,” dan “parafrasa,” menandakan bahwa *ChatGPT* dipandang sebagai alat bantu dalam penulisan dan kegiatan akademik.

5. KESIMPULAN

Penelitian ini mengkaji opini publik terhadap *ChatGPT* di pendidikan serta mengevaluasi performa algoritma *deep learning* dalam klasifikasi aspek dan sentimen. Dua aspek utama teridentifikasi melalui *LDA* yaitu *Asisten Belajar* dan *Etika Akademik*, dengan mayoritas sentimen positif (62,1% dan 55,3%). Model *Bi-LSTM* menunjukkan performa terbaik pada kedua tugas klasifikasi, khususnya pada split 90:10 dengan akurasi 91,11% (aspek) dan 87,88% (sentimen). Temuan ini mendukung pemanfaatan *Bi-LSTM* sebagai algoritma yang sesuai untuk analisis opini berbasis teks.

6. DAFTAR PUSTAKA

- [1] W. Suharmawan, "Pemanfaatan Chat GPT dalam dunia pendidikan," *Education Journal: Journal Educational Research and Development*, vol. 7, no. 2, pp. 158–166, 2023.
- [2] O. Sidoti, E. Park, and J. Gottfried, "About a quarter of U.S. teens have used ChatGPT for schoolwork – double the share in 2023," Pew Research Center, 2025. [Online]. Available: <https://pewrsr.ch/4g3Jqt0>
- [3] A. D. Pratama and H. Hendry, "Analisa Sentimen Masyarakat Terhadap Penggunaan CHATGPT Menggunakan Metode Support Vector Machine (SVM)," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 1, pp. 327–338, 2024. [Online]. Available: <http://www.jurnal.stkipggritulungagung.ac.id/index.php/jupi/article/view/4285>
- [4] M. G. Ramaputra and H. Purnomo, "Analisis sentimen opini masyarakat terhadap penggunaan ChatGPT di bidang pendidikan berbasis Twitter," *Jurnal Pepadun*, vol. 5, no. 3, pp. 275–285, 2024.
- [5] P. R. Amalia and E. Winarko, "Aspect-based sentiment analysis on Indonesian restaurant review using a combination of convolutional neural network and contextualized word embedding," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 3, pp. 285–294, 2021.
- [6] C. F. Hasri and D. Alita, "Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, pp. 145–160, 2022. [Online]. Available: <https://jim.teknokrat.ac.id/index.php/informatika/article/view/2026>
- [7] A. Maiti, A. Abarida, and M. Hanini, "The impact of feature extraction techniques on the performance of text data classification models," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 35, no. 2, pp. 1041–1052, 2024. [Online]. Available: <https://doi.org/10.11591/ijeecs.v35.i2.pp1041-1052>
- [8] D. G. Mandhasiya, H. Murfi, and A. Bustamam, "The hybrid of BERT and deep learning models for Indonesian sentiment analysis," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 1, pp. 591–602, 2024. [Online]. Available: <https://doi.org/10.11591/ijeecs.v33.i1.pp591-602>
- [9] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 12, no. 2, pp. 89–99, 2022.
- [10] D. Dharrao et al., "An efficient method for disaster tweets classification using gradient-based optimized convolutional neural networks with BERT embeddings," *MethodsX*, vol. 13, p. 102843, 2024. [Online]. Available: <https://doi.org/10.1016/j.mex.2024.102843>
- [11] N. N. Ainii, "Klasifikasi Sentimen, Topik, Dan Detail Topik Dari Ulasan Kai Access Menggunakan Multilayer Perceptron (MLP) Dan Bidirectional Long Short Term Memory (BiLSTM)," unpublished.
- [12] R. F. Hakim, "Simple CNN for Text Classification," *Medium*, 2024. [Online]. Available: <https://medium.com/@986110101/simple-cnn-for-text-classification-8ee92ce334f8>
- [13] G. A. Barnett, C. Calabrese, and J. B. Ruiz, "A comparison of three methods to determine the subject matter in textual data," *Frontiers in Research Metrics and Analytics*, vol. 8, p. 1104691, 2023.