

ANALISIS SENTIMEN PEMANFAATAN ARTIFICIAL INTELLIGENCE DALAM PENDIDIKAN DI INSTAGRAM DENGAN ALGORITMA RANDOM FOREST

Nirma Ayu Suryaningtyas¹, Devi Mustika², Ulum Hidayah Suryani³, Yusuf Nalim⁴
Sains Data, UIN K.H. Abdurrahman Wahid Pekalongan, Indonesia

Penulis korespondensi: tyasnirma1@gmail.com

ABSTRAK

Perkembangan teknologi di era industri 4.0, khususnya *Artificial Intelligence* (AI), telah membawa perubahan signifikan di berbagai bidang, termasuk pendidikan. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap pemanfaatan *Artificial Intelligence* dalam membantu tugas-tugas pelajar dan mahasiswa melalui komentar pada media sosial Instagram. Data dikumpulkan dengan teknik crawling dari tujuh reels Instagram terkait topik *Artificial Intelligence* dan pendidikan, menghasilkan 564 komentar. Analisis sentimen dilakukan menggunakan metode machine learning Random Forest, dengan tahapan preprocessing data seperti *Abbreviations Removal*; tokenisasi; *Stopword Removal*; dan lematisasi, lalu kemudian dilanjutkan dengan tahap pelabelan, analisis sentiment, dan visualisasi. Hasil penelitian menunjukkan bahwa sentimen netral mendominasi (68,4%), diikuti oleh sentimen negatif (8,6%) dan positif (22,8%). Evaluasi model Random Forest menunjukkan akurasi sebesar 67% dengan F1-Score rata-rata 38%. Temuan yang didapatkan dari penelitian ini diantaranya mayoritas masyarakat memiliki sentimen netral terhadap penggunaan *Artificial Intelligence* dalam Pendidikan, hal tersebut mencerminkan sikap hati-hati dan belum sepenuhnya menerima atau menolak. Meskipun ada apresiasi terhadap kemudahan yang ditawarkan *Artificial Intelligence*, banyak juga yang mengungkapkan kekhawatiran atas penyalahgunaannya oleh pelajar. Hasil wordcloud dan TF-IDF menunjukkan bahwa *Artificial Intelligence* paling sering digunakan untuk menyelesaikan tugas seperti esai. Hasil temuan tersebut menekankan pentingnya edukasi dan regulasi dalam penggunaan *Artificial Intelligence* agar manfaatnya maksimal dan risiko negatifnya dapat diminimalisir.

Kata kunci: Artificial Intelligence (AI), Analisis Sentimen, Random Forest, Pendidikan, Instagram

1 PENDAHULUAN

Perkembangan teknologi di industri 4.0 yang semakin pesat membawa berbagai kemudahan dalam menyelesaikan pekerjaan manusia dan memberikan dampak yang signifikan terhadap bidang kehidupan. Salah satu perkembangan teknologi di era industri 4.0 ini yaitu kecerdasan buatan atau *Artificial Intelligence* (AI) menjadi alat fenomenal dalam membuat pekerjaan manusia menjadi lebih efisien. Dalam beberapa tahun terakhir, penggunaan alat berbasis *Artificial Intelligence* mengalami peningkatan yang sangat signifikan. Fenomena ini menunjukkan tingginya minat masyarakat terhadap teknologi *Artificial Intelligence*, termasuk di Indonesia yang kini menjadi salah satu negara dengan tingkat keterlibatan tertinggi dalam pemanfaatan teknologi tersebut secara global (Helmiatin dkk., 2024). Hal tersebut tentunya membuktikan bahwa banyak masyarakat Indonesia yang mengikuti perkembangan teknologi *Artificial Intelligence* untuk

membantu memudahkan pekerjaan mereka. Salah satunya dibidang pendidikan, kemunculan *Artificial Intelligence* membawa perubahan yang signifikan dengan memberikan kemudahan dan inovasi baru seperti dengan adanya *Artificial Intelligence* chatbots atau pendeteksi plagiarisme (Rifky, 2024). Meskipun memiliki banyak manfaat, tidak selamanya perkembangan *Artificial Intelligence* mendapat persepsi atau sentimen positif dari masyarakat. Banyak masyarakat yang memberikan persepsi atau sentimen dengan berkomentar di seluruh media salah satunya media sosial. Sosial media dapat dimanfaatkan untuk memperoleh informasi opini (Putra & Wijaya, 2023). Salah satu media sosial tersebut adalah Instagram.

Instagram merupakan utilitas yang memungkinkan pengguna untuk mengekspresikan diri dengan memposting atau melihat foto, video, atau informasi di *Story Instagram*, Reels Instagram atau postingan biasa. Pengguna juga dapat menyampaikan komentar dengan berbagai karakter atau emoticon untuk mengekspresikan pendapat mereka terkait postingan tersebut. Komentar-komentar dari masyarakat dengan sentimen yang berbeda-beda tersebut jika dikumpulkan (*Crawling*) dapat dianalisis menjadi analisis sentimen lalu diproses menggunakan algoritma Random Forest, Random Forest (RA) merupakan pengembangan dari metode bagging yang termasuk dalam kelompok algoritma berbasis pohon (Tree) (Huda, 2023).

Berdasarkan latar belakang di atas, belum terdapat penelitian yang secara spesifik melakukan analisis sentimen terhadap pemanfaatan AI dalam pendidikan di Instagram dengan Algoritma Random Forest. Maka fokus pada penelitian ini untuk menganalisis sentimen dengan Algoritma Random Forest terkait pemanfaatan AI dalam pendidikan di Instagram. Hasil penelitian ini diharapkan dapat menjadi kontribusi terhadap penelitian lain terkait pemanfaatan AI dalam dunia pendidikan. Pembahasan dalam penelitian ini yaitu berisi analisis sentimen masyarakat terhadap perkembangan *Artificial Intelligence* yang digunakan dalam dunia pendidikan untuk membantu tugas serta pekerjaan sekolah atau kuliah pada media sosial Instagram. Analisis dilakukan dengan mengumpulkan komentar dari beberapa *Reels Instagram* yang membahas terkait penggunaan *Artificial Intelligence* dalam membantu tugas-tugas sekolah atau kuliah. Penelitian ini dilakukan dengan tujuan menyimpulkan opini masyarakat terkait penggunaan *Artificial Intelligence* dalam dunia pendidikan yang diunggah melalui media sosial Instagram untuk dikategorikan ke positif, netral hingga berbasis negatif dengan berbasis Machine Learning menggunakan metode Random Forest dan bahasa pemrograman Python.

2 METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode Random Forest untuk menganalisis sentimen penggunaan *Artificial Intelligence* pada pelajar atau mahasiswa di komentar Instagram. Algoritma Random Forest merupakan metode *ensemble learning* berbasis *decision tree*. Algoritma ini dikenal memiliki akurasi tinggi, mampu menangani data non-linear, serta dapat mengurangi risiko overfitting dibandingkan dengan metode lain seperti regresi linier (Fauzi, 2025). Dataset yang digunakan dalam penelitian ini terdiri dari 564 komentar yang merupakan hasil *crawling* di beberapa *Reels Instagram* yang membahas terkait penggunaan *Artificial Intelligence* untuk membantu dalam pendidikan atau membantu pelajar maupun mahasiswa. Adapun penelitian ini memiliki tahapan yang terdiri dari *Crawling Data*, *Globing Data*, *Labeling Text*, *Preprocessing Text*, *Analisis Sentimen (Random Forest)* dan *Visualization* seperti Gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Crawling Data

Crawling atau *Scraping Data* mengacu pada prosedur ekstraksi data yang dapat dilakukan secara otomatis dari situs web atau manual dengan cara *coding* menggunakan bahasa pemrograman *Python*. Tahap ini adalah proses yang sangat penting dalam bidang-bidang seperti Bisnis Intelijen di era modern. *Crawling* yang memungkinkan mengekstrak data terstruktur dari teks (*Text Mining*) seperti HTML, komentar atau *twets* dari Instagram maupun Twitter (Rizquina & Ratnasari, 2023). Dalam penelitian analisis sentimen ini, data dikumpulkan melalui teknik *crawling* pada tujuh komentar *reels* Instagram. Pencarian dilakukan dengan menggunakan kata kunci “Artificial Intelligence” dan “Pendidikan Indonesia” dalam rentang waktu 28 Maret 2025 hingga 3 Mei 2025. Proses pengambilan data menggunakan *web scraper*, dan mendapatkan tujuh *dataset* dalam format *Comma-Separated Values* (CSV). Data yang didapat secara keseluruhan sebanyak 564 komentar. Struktur *dataset* penelitian ini berupa data *frame* dengan dua atribut, yaitu atribut USERNAME yang merupakan nama pengguna dari akun yang memberikan komentar di *reels* Instagram dan COMMENT_TEXT yang merupakan atribut teks yang berisi komentar dari pengguna Instagram.

2.2 Globing Data

Data penelitian yang telah dikumpulkan akan di *glob* atau digabung agar lebih mudah diproses di tahap berikutnya. Dalam melakukan *globbing data*, atribut dalam file dataset harus sama, maksudnya dataset satu dengan yang lain sama-sama memiliki kolom komentar dan nama *user id*. Tahap penggabungan data komentar menjadi satu file *Comma-Separated Values* (CSV) dilakukan dengan menuliskan *syntax* menggunakan bahasa pemrograman *Python* pada aplikasi *Visual Studio Code*. Tahap ini penting dilakukan agar data lebih mudah diproses.

2.3 Preprocessing Text

Proses *text preprocessing* merupakan tahap krusial dalam *text mining*, berfungsi sebagai fondasi awal sebelum analisis lebih lanjut. Penelitian ini membahas tiga teknik utama *preprocessing* yang sangat signifikan untuk meningkatkan kualitas ekstraksi fitur teks sebelum pemodelan, diantaranya yaitu:

- a. Cleaning data atau proses penghapusan kata Stopword (*Stopword Removal*). Eliminasi kata-kata non-informatif (seperti "dan", "di", "yang"), tanda baca, spasi berlebih (*Abbreviations Removal*), dan penghapusan angka untuk mengurangi noise dalam data teks. Dalam preprocessing penelitian ini, data dibersihkan dari tanda baca, angka, spasi berlebih dan untuk memastikan input adalah string. Tujuan dari cleaning data ini adalah untuk mengurangi *noise* (Athira Luqyana dkk., 2018).

- b. Stemming merupakan proses reduksi kata ke bentuk dasarnya menggunakan algoritma *stemming* untuk menyatukan variasi morfologis. Data komentar yang telah melalui proses cleaning data akan menjadi bentuk kata dasar dengan diolah menggunakan library *sastrawi wordcloud matplotlib*. Proses tersebut disebut dengan stemming, yang merupakan prosedur komputasi yang mengurangi semua kata dengan akar kata yang sama (atau, jika pra perbaikan tidak disentuh, akar kata yang sama) ke bentuk yang umum umum, biasanya dengan menghilangkan setiap kata dari turunannya dan sufiks infleksi(Albab dkk., 2023).
- c. Pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*) Metode statistik untuk menilai pentingnya suatu kata dalam dokumen relatif terhadap seluruh korpus(Comtech dkk., 2025).

2.4 Labeling Text

Machine Learning telah digunakan untuk melakukan tugas-tugas di berbagai domain seperti klasifikasi, deteksi objek, segmentasi gambar, dan analisis bahasa alami. Proses pengklasifikasian ini disebut *Labeling Text* yang menjadi salah satu tugas terpenting dalam *machine learning* seperti analisis sentimen(Zhang dkk., 2021). Dalam analisis sentimen banyak sekali model yang dapat digunakan untuk *text labeling*, diantaranya yaitu model *TextBlob* dan *VADER*. Model *TextBlob* tersebut merupakan pendekatan berbasis leksikon yang bergantung pada asumsi bahwa orientasi semantik teks sangat terkait dengan polaritas kata dan frasa yang muncul di dalamnya(Catelli dkk., 2022). Penggunaan model ini memiliki keuntungan cepat dan mudah untuk diimplementasikan, tidak perlu data latih, serta dapat dikustomisasi dengan kata kunci spesifik seperti emoji dan slang(Nugraha, 2021). Pelabelan teks dilakukan untuk mengklasifikasi komentar yang masuk ke dalam kategori positif, netral maupun negatif. Kemudian model *VADER* (*Valence Aware Dictionary and Sentiment Reasoner*) merupakan alat analisis sentimen yang beroperasi pada leksikon dan pendekatan berbasis aturan. Alat ini bekerja dengan leksikon terlatih, yang terdiri dari daftar kata atau frasa yang sering digunakan untuk mengungkapkan pendapat atau perasaan tertentu(Leonardi dkk., 2025). Proses pelabelan data pada penelitian ini menggunakan model *Lexicon-Based TextBlob* dan *VADER*. Teknis dari model klasifikasi ini yaitu sistem *syntax* akan menentukan sentimen dari daftar kata positif dan negatif serta emoji untuk dihitung polaritas sentimennya menggunakan *Lexicon-Based (TextBlob)*(Abiola dkk., 2023). Kemudian, model *VADER* akan mengambil nilai sentimen numerik untuk dihitung *Compound*-nya. Nilai *polarity* dan *Compound* digabungkan dengan menghitung rata-rata sebagai skor gabungannya. Berikut merupakan *syntax* yang digunakan untuk pelabelan teks komentar.

```
# 3. PELABELAN SENTIMEN
def analyze_sentiment(text):
    # TextBlob Analysis
    polarity = TextBlob(text).sentiment.polarity

    # VADER Analysis
    analyzer = SentimentIntensityAnalyzer()
    vader_scores = analyzer.polarity_scores(text)

    # Gabungkan hasil
    if polarity > 0.2 or vader_scores['compound'] > 0.05:
        return 'positif'
    elif polarity < -0.2 or vader_scores['compound'] < -0.05:
        return 'negatif'
    else:
        return 'netral'
```

Gambar2. Hasil Labeling Text

2.5 Analisis Sentimen (Random Forest)

Dalam konteks analisis sentimen, metode *Random Forest* digunakan untuk mengklasifikasikan opini atau perasaan pengguna yang terkandung dalam data teks, seperti ulasan aplikasi, tweet, atau komentar, ke dalam kategori sentimen positif, negatif, atau netral (Afdhal dkk., 2022). Data yang sudah melalui tahap pre-processing dan pelabelan akan direpresentasikan numerik menggunakan metode seperti TF-IDF agar dapat diproses oleh algoritma (Herdiyani & Zailani, 2022). Atau dengan kata lain, *Random Forest* dalam penelitian ini digunakan sebagai model *machine learning* yang belajar dari representasi teks (TF-IDF) untuk klasifikasi sentimen. TF-IDF merupakan rumus yang digunakan untuk menghitung seberapa sering suatu kata muncul dalam suatu dokumen (TF tinggi) tapi jarang muncul dalam dokumen lain (IDF tinggi). Berikut persamaan dari pembobotan kata dengan menggunakan TF-IDF:

$$\text{IDF} = \log \frac{D}{DF} \quad (1)$$

atau

$$\text{TF-IDF} = \text{tf} \times \text{idf} \quad (2)$$

Keterangan:

D: Jumlah dokumen yang ada dalam data training

DF: Jumlah dokumen yang memiliki kata tersebut

Tf (term frequency): seberapa sering suatu kata muncul pada dokumen

idf (inverse document frequency): seberapa jarang kata tersebut muncul dalam Kumpulan dokumen. (Afdhal dkk., 2022).

Penerapan *Random Forest* dalam analisis sentimen memberikan pemahaman yang mendalam tentang persepsi publik terhadap produk, layanan, atau isu tertentu, serta membantu dalam pengambilan keputusan berbasis data yang lebih tepat dan responsif terhadap kebutuhan pengguna atau masyarakat. Dalam penelitian ini, model *Random Forest* juga digunakan sebagai prediksi dan evaluasi, memberikan laporan performa terkait akurasi, mencari parameter terbaik, dan mencari

distribusi kedalaman pohon. Berikut adalah *syntax* analisis sentimen menggunakan model *Random Forest*.

```
# 13. CEK AKURASI HASIL TERBAIK
param_grid = {
    'n_estimators': [50, 100, 200],
    'max_depth': [10, 20, None], # None artinya tidak ada batasan kedalaman
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4]
}

# A. Inisialisasi model Random Forest
rf = RandomForestClassifier(random_state=42, class_weight="balanced")
grid_search = GridSearchCV(estimator=rf, param_grid=param_grid, cv=5, scoring='accuracy', n_jobs=-1, verbose=2)

print("Memulai proses Grid Search...")
grid_search.fit(X_train, y_train)
print("Grid Search selesai.")

# B. Tampilkan parameter terbaik yang ditemukan
print("\nParameter Terbaik yang Ditemukan:")
print(grid_search.best_params_)

# C. Tampilkan skor terbaik pada data training (menggunakan cross-validation)
print("\nSkor Akurasi Terbaik (Cross-Validation pada Data Training):")
print(grid_search.best_score_)
```

Gambar 3. *Syntax Analisis Sentimen*

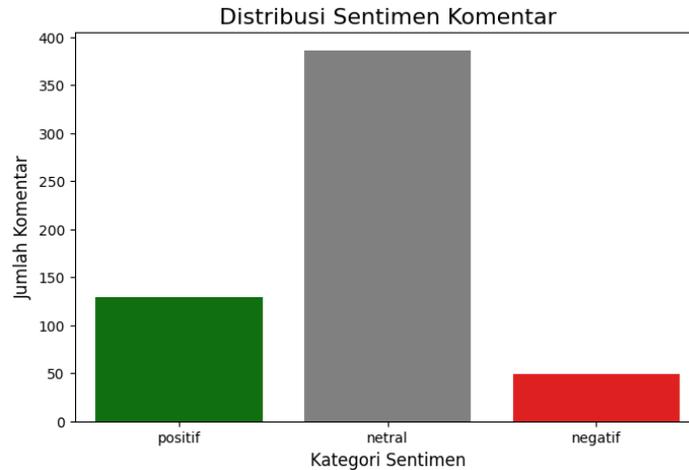
3 HASIL DAN PEMBAHASAN

3.1 Analisis Distribusi Sentimen

Data dalam penelitian ini dianalisis menggunakan bahasa pemrograman Python dan aplikasi Google Colab, dengan memproses sebanyak 564 komentar yang diambil dari tujuh reels Instagram terkait topik pemanfaatan *Artificial Intelligence* (AI) dalam mendukung proses pembelajaran maupun penyelesaian tugas oleh pelajar dan mahasiswa. Sentimen yang dianalisis dikategorikan ke dalam tiga kelas, yaitu:

1. Sentimen positif (129 ulasan atau 22,8%), yang mencerminkan pandangan mengenai manfaat, kepuasan, serta apresiasi terhadap hadirnya kecerdasan buatan.
2. Sentimen netral (386 ulasan atau 68,4%), yang berisi opini pengguna yang tidak secara tegas menunjukkan sikap positif maupun negatif, serta pandangan yang mendukung penggunaan *Artificial Intelligence* namun menekankan pentingnya penggunaan yang tidak berlebihan.
3. Sentimen negatif (49 ulasan atau 8,6%), yang memuat kekhawatiran terhadap kecenderungan pelajar atau mahasiswa meminta referensi atau bantuan dari *Artificial Intelligence* secara mentah-mentah, serta komentar bernada sarkastik atau kritik terkait dampak negatif penggunaan AI dalam dunia akademik.

Berikut interpretasi dari hasil distribusi sentimen pemanfaatan *Artificial Intelligence* (AI) dalam mendukung proses pembelajaran maupun penyelesaian tugas oleh pelajar dan mahasiswa.

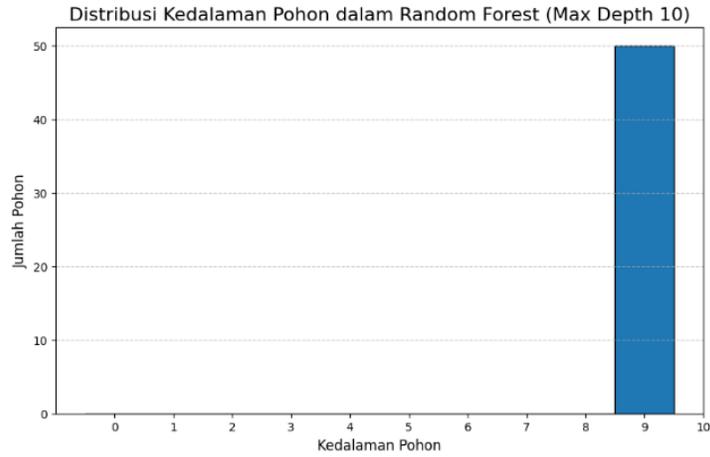


Gambar 4. Hasil Distribusi Analisis Sentimen

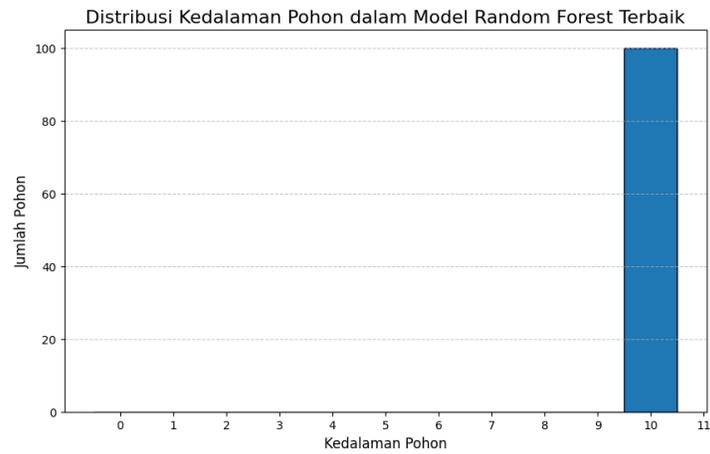
3.2 Evaluasi Model Klasifikasi

Pengujian paramater pada penelitian ini membatasi kedalaman *tree* sehingga semuanya menjadi 10 (konsisten) dan jumlah pohon sebanyak 50, hal tersebut dilakukan agar model lebih generalis, mudah diprediksi dan diinterpretasi serta mengurangi risiko overfitting dibanding model dengan kedalaman >100 . Selain itu, hal tersebut juga dikarenakan dataset yang kecil yakni hanya 113 sampel. Pada pengujian parameter penelitian ini, menerapkan SMOTE untuk meningkatkan *recall* negatif dan positif. Hasil pengujian dengan parameter terbaik berada pada kedalaman 10 dengan rata-rata tanpa bobot (Macro Avg) yaitu *precision* sebesar 42%, *recall* 39%, F1-Score 38% dan nilai akurasi sebesar 67% yang artinya akurasi cukup tinggi (67%), tapi macro precision, recall, dan F1-score rendah. Hal tersebut menunjukkan model hanya cenderung bagus pada kelas mayoritas.

Pada pegujian parameter penelitian ini, kedalaman 10 menghasilkan nilai *precision* kelas negatif 0% dari yang diprediksi benar, netral 73% artinya banyak prediksi netral yang sebenarnya bukan netral, dan *precision* positif 54%. Berdasarkan nilai *precision* tersebut, membuktikan bahwa tingkat kemampuan sistem dalam mencari ketetapan informasi prediksi kurang baik karena terjadi ketidakseimbangan kelas atau terlalu sering memprediksi kelas netral (*precision* rendah). Lalu, kemampuan sistem dalam mengenali sentimen atau *recall* pada kelas negatif 0% artinya sulit untuk dideteksi dan positif 26% yang terdeteksi. Sedangkan, untuk kelas netral model sangat baik dalam mendeteksi yaitu sebesar 90%. Perbandingan rata-rata antara *precision* dan *recall* disebut *F1-Score*, pada kelas negatif skornya sebesar 0%, sedangkan untuk kelas netral 80% dan positif 35%. Berdasarkan penjelasan tersebut, dapat disimpulkan nilai *F1-Score* kelas netral sangat baik, karena kemampuan sistem dalam memprediksi kelas netral tergolong tinggi. Penjelasan diatas dapat dilihat dan dipahami melalui interpretasi gambar dan tabel dibawah ini.



Gambar 5. Diagram Distribusi Kedalaman Pohon



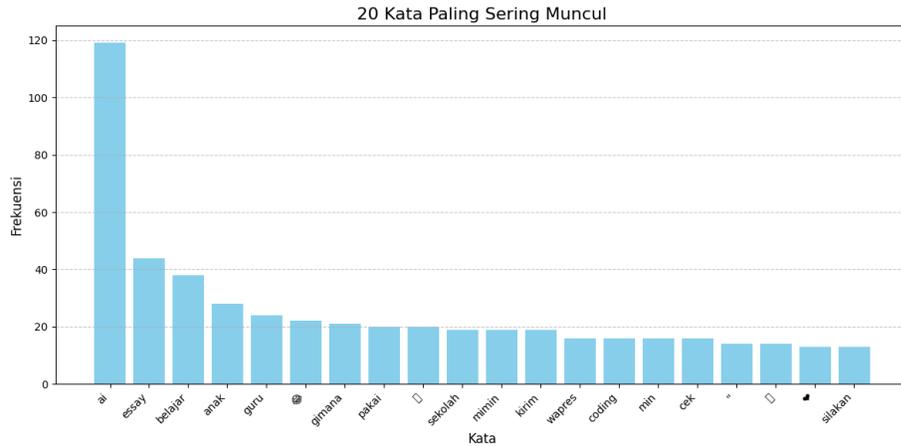
Gambar 6. Distribusi Kedalaman Pohon Terbaik

	precision	recall	f1-score	support
Negative	0.00	0.00	0.00	9
Neutral	0.73	0.90	0.80	77
Positive	0.54	0.26	0.35	27
accuracy			0.67	113
macro avg	0.42	0.39	0.38	113
Weighted avg	0.62	0.67	0.63	113
Akurasi	0.672566371681416			

Tabel 1. Hasil Klasifikasi Parameter Sentimen

3.3 Analisis WorldCloud atau Pembobotan Kata

Pada penelitian ini dilakukan pengujian menggunakan *Worldcloud* untuk mengetahui frekuensi kata atau topik yang sering diperbincangkan. Berikut merupakan interpretasi dari frekuensi kata yang sering muncul.



Gambar 7. Diagram Frekuensi Kata



Gambar 8. Visualisasi WorldCloud

Berdasarkan gambar 7 dan 8, didapatkan kata yang paling sering muncul atau menjadi topik perbincangan dalam kolom komentar adalah “ai” atau *Artificial Intelligence* dengan frekuensi sebanyak 119 kali. Selain visualisasi *Wordcloud*, pada pengujian ini juga dilakukan pembobotan kata dengan bobot TF-IDF tertinggi untuk menemukan kata unik atau mengetahui seberapa sering suatu kata muncul pada dokumen tertentu tapi jarang muncul di dokumen lain (umum). Hasil dari TF-IDF akan berbeda dengan hasil analisis kata dengan frekuensi kemunculan yang tinggi, hal itu dikarenakan analisis frekuensi kemunculan kata tertentu mencakup semua dokumen atau dalam penelitian ini yaitu berupa komentar, sedangkan kata yang sering muncul dalam dokumen tertentu disebut kata unik atau kata yang memiliki bobot TF-IDF tinggi. Hasil dari TF-IDF dalam penelitian ini, diperoleh 2 kata yang paling sering muncul pada dokumen tertentu juga merujuk kedalam konteks penelitian, diantaranya kata “essay” dengan skor 38.83 serta kata “ai” dengan skor 22.27. Interpretasi dari pembobotan TF-IDF dapat dilihat pada tabel dibawah ini.

Kata dengan Bobot TF-IDF Tertinggi	
Term	Rank
essay	38.832433
ai	22.276299
min	12.622332
kirim	12.268784
mimin	12.268784

Tabel 2. Hasil Pembobotan Kata

Berdasarkan tabel 2 di atas diperoleh implikasi untuk model bahwa kata “essay” relevan dengan tugas akademik yang kemungkinan berupa soal uraian atau karya tulis ilmiah dan kata “ai” merujuk pada singkatan dari *Artificial Intelligence*, hal tersebut dapat disimpulkan banyak pelajar atau mahasiswa yang memanfaatkan teknologi *Artificial Intelligence* untuk membantu tugas akademik mereka. Penerapan *Artificial Intelligence* dalam pendidikan mengubah peran pendidik. Pendidik harus beradaptasi dengan perubahan ini dan mengembangkan keterampilan baru dalam menggunakan teknologi AI. Meskipun AI dapat membantu dalam personalisasi pembelajaran dan evaluasi, peran pendidik sebagai fasilitator, motivator, dan pendamping pelajar tetap penting untuk memberikan pengalaman belajar yang holistik dan mendukung perkembangan pelajar secara menyeluruh(Rifky, 2024).

4 KESIMPULAN

Berdasarkan hasil analisis sentimen terhadap 564 komentar Instagram mengenai penggunaan *Artificial Intelligence* dalam dunia pendidikan, dapat disimpulkan bahwa persepsi masyarakat cenderung netral, dengan sebagian besar komentar menekankan pentingnya penggunaan *Artificial Intelligence* secara proporsional, contohnya komentar dari akun @amellia.dae yang menyatakan “manfaat in buat ningkatin performa dan mengembangka ide ga berarti ketergantungan juga, mikir iya, nyatet iya, ngerekam iya, ai iya”, komentar dari akun @mauziahzs yaitu “AI dibuat untuk membantu aktivitas manusia, intinya jangan sampai kecanduan aja, dikit-dikit AI”, dan komentar

dari akun @choiieunw yaitu “bukan berarti kita harus pakai Ai terus terusan yha gess itu juga ngga baik, cuman aku ngasih info aja kalau pakai Ai itu ngga asal tanya tpi ada kata kunci untuk mencari materi supaya lebih spesifik dan mungkin materi nya lebih lengkap dan mendalam”. Meskipun terdapat apresiasi terhadap manfaat AI dalam memudahkan tugas akademik, kekhawatiran terhadap potensi penyalahgunaan dan dampak negatif juga cukup tinggi. Model Random Forest yang digunakan dalam penelitian ini mampu mengklasifikasikan sentimen dengan akurasi yang cukup baik, yaitu 67%. Hasil wordcloud dan pembobotan TF-IDF menunjukkan bahwa topik terkait “essay” dan “ai” paling sering diperbincangkan, menandakan bahwa *Artificial Intelligence* banyak dimanfaatkan untuk membantu tugas-tugas akademik. Penelitian ini merekomendasikan perlunya edukasi dan regulasi terkait penggunaan *Artificial Intelligence* di lingkungan pendidikan agar manfaat yang diperoleh dapat optimal dan risiko penyalahgunaan dapat diminimalisir.

Melalui hasil penelitian ini, disarankan agar riset selanjutnya dapat memperluas sumber data, misalnya dengan mengambil komentar dari media sosial lain seperti Twitter atau TikTok untuk mendapatkan sudut pandang yang lebih beragam. Selain itu, akan lebih baik jika model algoritma yang digunakan tidak hanya terbatas pada Random Forest, tetapi juga dibandingkan dengan algoritma lain seperti SVM, Naive Bayes, atau bahkan pendekatan deep learning seperti LSTM, agar dapat diketahui mana yang paling efektif dalam analisis sentimen. Pengembangan metode pelabelan teks juga penting dilakukan, misalnya dengan pendekatan *semi-supervised* atau *active learning*, supaya hasil klasifikasinya lebih akurat dan efisien, terutama saat diterapkan pada data yang terus berkembang. Ke depan, hasil analisis sentimen ini juga bisa dijadikan dasar untuk membangun sistem pendeteksi dini guna mengantisipasi penyalahgunaan teknologi AI oleh pelajar.

UCAPAN TERIMAKASIH

Penulis mengucapkan terimakasih sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dalam penyusunan artikel ini. Terutama kepada bapak Yusuf Nalim sebagai pembimbing atas masukan dan bimbingan yang telah diberikan dalam proses penelitian dan penyusunan artikel. Penulis juga mengucapkan terimakasih kepada keluarga, rekan sejawat dan semua pihak yang tidak disebutkan satu persatu yang telah memberikan dukungan dan doa selama penyusunan artikel ini. Ucapan terimakasih disampaikan dalam satu paragraf singkat. Ucapan terimakasih dapat ditujukan pada pihak-pihak yang telah membantu dalam penelitian, baik dalam bentuk support dana, perizinan, konsultan, maupun membantu dalam pengambilan data.

DAFTAR PUSTAKA

- Abiola, O., Abayomi-Alli, A., Tale, O. A., Misra, S., & Abayomi-Alli, O. (2023). Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser. *Journal of Electrical Systems and Information Technology*, 10(1).
<https://doi.org/10.1186/s43067-023-00070-9>
- Afdhal, I., Kurniawan, R., Iskandar, I., Salambue, R., Budianita, E., & Syafria, F. (2022). Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia. *Jurnal Nasional Komputasi dan Teknologi Informasi*, 5(1), 122–130.
<http://ojs.serambimekkah.ac.id/jnkti/article/view/4004/pdf>
- Albab, M. U., P., Y. K., & Fawaiq, M. N. (2023). Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic. *Jurnal Transformatika*, 20(2), 1–12.
<https://doi.org/10.26623/transformatika.v20i2.5374>
- Athira Luqyana, W., Cholissodin, I., & Perdana, R. S. (2018). Analisis Sentimen Cyberbullying

- pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), 4704–4713. <http://j-ptiik.ub.ac.id>
- Catelli, R., Pelosi, S., & Esposito, M. (2022). Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. *Electronics (Switzerland)*, 11(3). <https://doi.org/10.3390/electronics11030374>
- Comtech, J. I., Darmawan, I., Arifin, M. N., Ramadhani, N., Puspa, N., & Informatika, P. S. (2025). *ALGORITMA OKAPI BM25 DALAM SISTEM PENCARIAN*. 10(1).
- Fauzi, A. (2025). *Journal of Artificial Intelligence and Digital Business (RIGGS) Prediksi Harga Properti Di Indonesia Menggunakan Algoritma Random*. 4(1), 43–49.
- Helmiatin, Hidayat, A., & Kahar, M. R. (2024). Investigating the adoption of AI in higher education: a study of public universities in Indonesia. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2380175>
- Herdiyani, T. C., & Zailani, A. U. (2022). Sentiment Analysis Terkait Pemindahan Ibu Kota Indonesia Menggunakan Metode Random Forest Berdasarkan Tweet Warga Negara Indonesia. *Jurnal Teknologi Sistem Informasi*, 3(2), 154–165. <https://doi.org/10.35957/jtsi.v3i2.2920>
- Huda, D. N. I. (2023). Analisis Sentimen Perbandingan Layanan Jasa Pengiriman Kurir Pada Ulasan Play Store Menggunakan Metode Random Forest dan Decision Tree. *Jurnal Ilmiah Informatika*, 11(2), 150–158.
- Leonardi, V. H., Ibrahim, A., Kurnia, R. D., & Afrina, M. (2025). Analisis Sentimen Ulasan Aplikasi Pembelajaran Bahasa Menggunakan Metode VADER. *Jurnal Algoritma*, 22(1), 767–776. <https://doi.org/10.33364/algoritma/v.22-1.2285>
- Nugraha, K. A. (2021). Analisis Sentimen Berbasis Emoticon pada Komentar Instagram Bahasa Indonesia Menggunakan Naïve Bayes. *Jurnal Teknik Informatika dan Sistem Informasi*, 7(3), 715–721. <https://doi.org/10.28932/jutisi.v7i3.4094>
- Putra, S. A., & Wijaya, A. (2023). ANALISIS SENTIMEN ARTIFICIAL INTELLIGENCE (AI) PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE LEXICON BASED. *JuSiTik : Jurnal Sistem dan Teknologi Informasi Komunikasi*, 7(1). <https://doi.org/10.32524/jusitik.v7i1.1042>
- Rifky, S. (2024). Dampak Penggunaan Artificial Intelligence Bagi Pendidikan Tinggi. *Indonesian Journal of Multidisciplinary on Social and Technology*, 2(1), 37–42. <https://doi.org/10.31004/ijmst.v2i1.287>
- Rizquina, A. Z., & Ratnasari, C. I. (2023). Implementasi Web Scraping untuk Pengambilan Data Pada Website E-Commerce. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(4), 377–383. <https://doi.org/10.47233/jteksis.v5i4.913>
- Zhang, S., Jafari, O., & Nagarkar, P. (2021). *A Survey on Machine Learning Techniques for Auto Labeling of Video, Audio, and Text Data*. 1–13. <http://arxiv.org/abs/2109.03784>