



# Identifikasi Indikasi Risiko Depresi pada Unggahan Media Sosial X Menggunakan Natural Language Processing dan Algoritma Random Forest

Arif Siswandi\*, Arif Susilo

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia

Email: <sup>1,\*</sup>arif.siswandi@pelitabangsa.ac.id, <sup>2</sup>arif.susilo@pelitabangsa.ac.id

Email Penulis Korespondensi: nurhadi@pelitabangsa.ac.id

**Abstrak**—Depresi pada mahasiswa merupakan isu kesehatan mental yang semakin mendapat perhatian karena dapat memengaruhi kualitas hidup dan performa akademik. Media sosial X sebagai platform berbasis teks menyediakan ruang ekspresi spontan yang berpotensi mencerminkan kondisi emosional pengguna. Penelitian ini bertujuan menganalisis pola linguistik yang diasosiasikan dengan indikasi ekspresi depresi pada unggahan media sosial menggunakan pendekatan Natural Language Processing (NLP) dan algoritma Random Forest. Data dikumpulkan melalui teknik scraping pada periode Januari hingga November 2024 menggunakan kata kunci yang secara konseptual mengacu pada indikator Patient Health Questionnaire-9 (PHQ-9) dan disesuaikan dengan representasi bahasa yang umum digunakan pada media sosial. Dari 36.081 unggahan yang diperoleh, dilakukan proses seleksi melalui penghapusan duplikasi, penyaringan bahasa, dan eliminasi konten tidak relevan sehingga diperoleh 1.070 data yang digunakan sebagai dataset penelitian. Tingkat penyaringan yang tinggi menunjukkan bahwa banyak unggahan hasil scraping tidak secara langsung merepresentasikan ekspresi emosional yang relevan. Dataset kemudian diberi label secara manual ke dalam tiga kategori indikasi depresi, yaitu ringan, sedang, dan berat. Tahapan analisis meliputi preprocessing teks, ekstraksi fitur TF-IDF, serta pemodelan klasifikasi Random Forest. Hasil evaluasi menunjukkan akurasi sebesar 97%, namun nilai ini perlu ditafsirkan secara hati-hati karena performa model dapat dipengaruhi oleh karakteristik dataset dan proses pelabelan manual. Oleh karena itu, model lebih tepat dipahami sebagai pendekatan eksploratif untuk mengidentifikasi pola bahasa yang berkaitan dengan ekspresi emosional pada teks media sosial, bukan sebagai alat diagnosis klinis.

**Kata Kunci:** Natural Language Processing; Analisis Teks Media Sosial; Ekspresi Depresi; Machine Learning; Random Forest

**Abstract**—Depression among university students has become an important mental health concern due to its potential impact on quality of life and academic performance. Social media platform X, as a text-based communication medium, provides a space for spontaneous expression that may reflect users' emotional states. This study aims to analyze linguistic patterns associated with indicative depressive expressions in social media posts using a Natural Language Processing (NLP) approach and the Random Forest algorithm. Data were collected through web scraping between January and November 2024 using keywords conceptually derived from the Patient Health Questionnaire-9 (PHQ-9) indicators and adapted to linguistic expressions commonly used in social media communication. From an initial collection of 36,081 posts, several filtering stages were conducted, including duplicate removal, language filtering, and elimination of irrelevant content, resulting in a final dataset of 1,070 posts used in this study. The high filtering rate indicates that many scraped posts did not directly represent relevant emotional expressions. The dataset was manually labeled into three indicative categories of depressive expressions: mild, moderate, and severe. The analytical process included text preprocessing, TF-IDF feature extraction, and classification modeling using the Random Forest algorithm. The evaluation results show an accuracy of 97%. However, this value should be interpreted cautiously because model performance may be influenced by dataset characteristics and the manual labeling process. Therefore, the proposed model should be regarded as an exploratory approach for identifying linguistic patterns associated with emotional expressions in social media text rather than a clinical diagnostic tool for depression.

**Keywords:** Natural Language Processing; Social Media Text Analysis; Depressive Expressions; Machine Learning; Random Forest

## 1. PENDAHULUAN

Isu kesehatan mental, khususnya depresi [1], semakin menjadi perhatian global karena berkaitan dengan meningkatnya risiko bunuh diri, penurunan kualitas hidup, serta gangguan psikologis jangka panjang [2]. Kelompok mahasiswa merupakan populasi yang rentan terhadap gangguan kesehatan mental akibat tekanan akademik [3], tuntutan sosial, serta ketidakpastian masa depan yang dihadapi pada fase transisi menuju kedewasaan [4]. Kondisi tersebut sering kali memunculkan stres berkepanjangan, kelelahan emosional, serta perasaan tidak berdaya yang dapat berkembang menjadi gejala depresi apabila tidak terdeteksi sejak dini [4]. Namun demikian, sebagian mahasiswa tidak menyadari gejala awal yang dialami atau enggan mencari bantuan profesional karena adanya stigma sosial terhadap gangguan kesehatan mental.

Perkembangan teknologi digital telah mengubah pola komunikasi generasi muda, terutama melalui media sosial. Platform media sosial seperti X (Twitter) [5] menyediakan ruang ekspresi yang relatif bebas dan spontan bagi pengguna untuk menyampaikan pengalaman personal, emosi, serta tekanan psikologis yang dialami [6]. Unggahan teks pada media sosial tidak hanya berfungsi sebagai sarana komunikasi sosial, tetapi juga dapat merefleksikan kondisi emosional individu melalui pilihan kata, pola ekspresi, serta intensitas bahasa yang digunakan [7]. Dengan demikian, media sosial memiliki potensi sebagai sumber data alternatif untuk memahami dinamika kesehatan mental pengguna secara tidak langsung melalui analisis pola linguistik dalam teks [8].

Seiring berkembangnya penelitian di bidang Natural Language Processing (NLP) [9], berbagai pendekatan komputasional telah digunakan untuk menganalisis indikasi depresi pada teks media sosial [10]. Penelitian sebelumnya menunjukkan bahwa metode machine learning klasik seperti Support Vector Machine, Naïve Bayes, dan Random Forest mampu mengklasifikasikan emosi atau gejala depresi dalam teks dengan performa yang cukup baik [11]. Selain itu,



perkembangan terbaru juga menunjukkan bahwa model deep learning [12] berbasis transformer, seperti BERT atau XLNet, mampu menangkap konteks semantik yang lebih kompleks dalam teks dan sering menghasilkan performa klasifikasi yang lebih tinggi. Meskipun demikian, penerapan model deep learning tidak selalu menjadi pilihan optimal dalam semua konteks penelitian, terutama ketika dataset relatif terbatas, bahasa yang digunakan sangat informal, atau sumber daya komputasi terbatas. Dalam kondisi tersebut, algoritma machine learning klasik yang lebih interpretable dan efisien tetap relevan untuk dikembangkan sebagai pendekatan alternatif dalam analisis teks kesehatan mental.

Meskipun berbagai penelitian telah mengeksplorasi deteksi depresi berbasis media sosial [1], sebagian besar penelitian masih berfokus pada klasifikasi sentimen umum (positif, negatif, netral) atau deteksi depresi secara biner [13]. Pendekatan tersebut belum sepenuhnya mampu menggambarkan tingkat keparahan gejala depresi secara bertahap, yang sebenarnya lebih relevan dalam konteks pemantauan kesehatan mental. Selain itu, penelitian yang secara khusus mengkaji pola linguistik mahasiswa Indonesia pada media sosial masih relatif terbatas, terutama terkait penggunaan bahasa informal, slang, serta variasi ekspresi emosional yang khas pada kelompok usia tersebut [9]. Kondisi ini menunjukkan adanya celah penelitian dalam pengembangan model klasifikasi teks yang mampu mengidentifikasi indikasi risiko depresi secara multikelas pada unggahan mahasiswa berbahasa Indonesia [14].

Dalam penelitian ini, indikator gejala depresi yang diadaptasi dari Patient Health Questionnaire-9 (PHQ-9) digunakan sebagai kerangka konseptual untuk menyusun kata kunci awal dan pedoman pelabelan data [15]. Penting untuk ditegaskan bahwa PHQ-9 dalam penelitian ini tidak digunakan sebagai alat diagnosis klinis, melainkan sebagai referensi konseptual untuk mengidentifikasi representasi linguistik dari gejala depresi dalam teks [16]. Dengan demikian, pendekatan yang digunakan bertujuan untuk mengidentifikasi indikasi risiko depresi berbasis pola bahasa, bukan untuk menggantikan asesmen psikologis profesional [17].

Berdasarkan celah penelitian tersebut, studi ini mengembangkan model klasifikasi teks untuk mengidentifikasi indikasi tingkat risiko depresi pada unggahan mahasiswa di media sosial X menggunakan pendekatan Natural Language Processing [18]. Representasi fitur teks diperoleh melalui metode Term Frequency–Inverse Document Frequency (TF-IDF) dengan skema n-gram, sedangkan proses klasifikasi dilakukan menggunakan algoritma Random Forest yang dikenal memiliki kemampuan baik dalam menangani data berdimensi tinggi serta distribusi kelas yang tidak seimbang [19].

Kontribusi utama penelitian ini terletak pada pengembangan kerangka klasifikasi multikelas berbasis representasi linguistik gejala depresi pada unggahan mahasiswa berbahasa Indonesia. Pendekatan ini berupaya memperluas penelitian sebelumnya yang umumnya terbatas pada klasifikasi sentimen atau deteksi depresi biner, dengan memetakan indikasi tingkat keparahan gejala secara bertahap melalui analisis teks media sosial [20]. Berdasarkan latar belakang tersebut, tujuan utama penelitian ini adalah mengembangkan dan mengevaluasi model klasifikasi teks berbasis Natural Language Processing untuk mengidentifikasi indikasi tingkat risiko depresi pada unggahan mahasiswa di media sosial X [5], serta menganalisis performa algoritma Random Forest dalam mengklasifikasikan tingkat indikasi tersebut secara multikelas berdasarkan pola linguistik yang muncul dalam teks [21].

## 2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan text mining berbasis Natural Language Processing (NLP) dan supervised machine learning untuk mengidentifikasi indikasi depresi pada unggahan pengguna media sosial X.



**Gambar 1.** Tahapan Penelitian

### 2.1. Pengumpulan Data

Data penelitian diperoleh dari unggahan pengguna media sosial X (Twitter) menggunakan teknik web scraping melalui Application Programming Interface (API). Pengambilan data dilakukan pada periode Januari 2024 hingga November 2024 dengan filter bahasa Indonesia.

Kata kunci pencarian disusun berdasarkan indikator gejala depresi dari instrumen Patient Health Questionnaire-9 (PHQ-9) yang kemudian diterjemahkan ke dalam representasi linguistik yang sering muncul pada unggahan media sosial. Dataset awal yang diperoleh kemudian melalui proses seleksi untuk memastikan kualitas data yang digunakan dalam penelitian.

**Tabel 1.** Proses Seleksi Dataset

Tahap	Jumlah Data
Hasil scraping awal	36.081
Penghapusan duplikasi	4.215
Penyaringan bahasa	18.562
Penghapusan konten tidak relevan	12.234
Dataset akhir	1.070



Tabel 1 menunjukkan proses seleksi dataset yang diperoleh dari hasil scraping media sosial X. Dari total 36.081 unggahan awal, dilakukan beberapa tahap penyaringan yaitu penghapusan data duplikat, penyaringan bahasa untuk memastikan kesesuaian dengan Bahasa Indonesia, serta penghapusan konten yang tidak relevan dengan konteks penelitian. Setelah melalui proses pembersihan data tersebut, diperoleh 1.070 unggahan yang digunakan sebagai dataset akhir dalam penelitian ini.

## 2.2. Adaptasi Indikator PHQ-9

Indikator gejala depresi pada PHQ-9 digunakan sebagai kerangka konseptual untuk mengidentifikasi ekspresi linguistik yang berkaitan dengan kondisi emosional pengguna media sosial.

**Tabel 2.** Adaptasi Indikator PHQ-9 ke Representasi Linguistik

Indikator PHQ-9	Deskripsi Gejala	Representasi Linguistik
Kehilangan minat	Tidak tertarik melakukan aktivitas	tidak semangat, malas, tidak mood
Perasaan sedih	Perasaan sedih berkepanjangan	sedih, hampa, kosong
Gangguan tidur	Kesulitan tidur	susah tidur, insomnia
Kelelahan	Merasa lelah hampir setiap hari	capek hidup, lelah banget
Perasaan tidak berharga	Merasa gagal atau tidak berguna	gagal, tidak berguna
Sulit konsentrasi	Kesulitan fokus	tidak fokus, pikiran kosong
Perubahan nafsu makan	Nafsu makan berkurang	tidak nafsu makan
Putus asa	Kehilangan harapan	menyerah, putus asa
Pikiran kematian	Keinginan untuk mati	ingin mati, bunuh diri

Tabel 2 menunjukkan proses adaptasi indikator gejala depresi dari instrumen Patient Health Questionnaire-9 (PHQ-9) ke dalam bentuk representasi linguistik yang umum digunakan dalam unggahan media sosial. Setiap indikator PHQ-9 diterjemahkan menjadi kata atau frasa yang merepresentasikan ekspresi emosional pengguna, sehingga dapat digunakan sebagai dasar dalam proses pengumpulan data, pelabelan dataset, serta analisis teks menggunakan pendekatan Natural Language Processing (NLP). Adaptasi ini bertujuan untuk menjembatani konsep gejala depresi secara klinis dengan pola bahasa yang muncul dalam data media sosial.

## 2.3. Pelabelan Data

Dataset yang telah diseleksi kemudian diberi label secara manual berdasarkan tingkat indikasi depresi yang muncul dalam teks.

**Tabel 3.** Kategori Label Dataset

Label	Kategori	Deskripsi
0	Depresi ringan	Ekspresi emosi negatif ringan
1	Depresi sedang	Ekspresi tekanan psikologis
2	Depresi berat	Ekspresi keputusan atau ide bunuh diri

Tabel 3 menunjukkan kategori label yang digunakan dalam proses pelabelan dataset pada penelitian ini. Setiap unggahan teks diklasifikasikan ke dalam tiga tingkat indikasi depresi, yaitu depresi ringan, depresi sedang, dan depresi berat, yang kemudian direpresentasikan dalam bentuk label numerik (0, 1, dan 2). Pelabelan ini dilakukan untuk menyediakan ground truth dalam proses supervised machine learning sehingga model dapat mempelajari pola linguistik yang berkaitan dengan tingkat indikasi depresi pada teks media sosial. Distribusi dataset pada setiap kategori ditunjukkan pada Tabel 4 berikut:

**Tabel 4.** Distribusi Dataset

Kategori	Jumlah Data
Depresi ringan	420
Depresi sedang	365
Depresi berat	285
Total	1.070

Tabel 4 menunjukkan distribusi jumlah data pada setiap kategori tingkat depresi dalam dataset yang digunakan pada penelitian ini. Dataset terdiri dari 1.070 unggahan teks yang telah melalui proses seleksi dan pelabelan, dengan distribusi 420 data pada kategori depresi ringan, 365 data pada kategori depresi sedang, dan 285 data pada kategori depresi berat. Distribusi ini menunjukkan bahwa dataset memiliki variasi representasi pada setiap kategori yang digunakan dalam proses pelatihan dan evaluasi model klasifikasi.

## 2.4. Preprocessing Data

Preprocessing dilakukan untuk membersihkan dan menstandarkan teks media sosial dengan menghapus elemen tidak relevan serta menormalisasi kata sebelum ekstraksi fitur dan pemodelan machine learning.



**Tabel 5.** Tahapan Preprocessing

Tahap	Deskripsi
Case Folding	Mengubah seluruh huruf menjadi huruf kecil
Cleaning	Menghapus URL, simbol, angka, dan tanda baca
Slang Normalization	Mengubah kata tidak baku menjadi bentuk formal
Tokenizing	Memecah kalimat menjadi kata
Stopword Removal	Menghapus kata umum yang tidak informatif
Stemming	Mengubah kata menjadi bentuk dasar

Tabel 5 menjelaskan tahapan preprocessing yang dilakukan pada data teks sebelum proses analisis menggunakan metode machine learning. Preprocessing bertujuan untuk membersihkan dan menormalisasi teks sehingga data menjadi lebih terstruktur dan siap diproses pada tahap ekstraksi fitur.

**Tabel 6.** Contoh Preprocessing Teks

Tahap	Contoh
Teks asli	Aku capek banget hidup begini rasanya pengen menyerah
Case folding	aku capek banget hidup begini rasanya pengen menyerah
Normalisasi	aku lelah sekali hidup begini rasanya ingin menyerah
Tokenizing	[aku, lelah, sekali, hidup, begini, rasanya, ingin, menyerah]
Stopword removal	[lelah, hidup, rasanya, menyerah]
Stemming	[lelah, hidup, rasa, serah]

Tabel 6 menunjukkan contoh proses preprocessing teks yang dilakukan pada data unggahan media sosial sebelum tahap analisis menggunakan metode machine learning. Tahapan preprocessing bertujuan untuk membersihkan dan menstandarkan teks sehingga lebih mudah diproses oleh model. Proses ini dimulai dari teks asli, kemudian dilakukan case folding untuk menyeragamkan huruf menjadi huruf kecil, diikuti normalisasi kata tidak baku agar sesuai dengan bentuk bahasa Indonesia formal.

Selanjutnya dilakukan tokenizing untuk memecah kalimat menjadi unit kata, stopwords removal untuk menghapus kata umum yang tidak memiliki kontribusi signifikan terhadap analisis, serta stemming untuk mengubah kata ke bentuk dasarnya. Tahapan ini menghasilkan representasi teks yang lebih terstruktur dan siap digunakan dalam proses ekstraksi fitur dan pemodelan klasifikasi.

## 2.5. Ekstraksi Fitur TF-IDF

Representasi numerik teks diperoleh menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) yang memberikan bobot pada kata berdasarkan frekuensi kemunculannya dalam dokumen.

**Tabel 7.** Parameter TF-IDF

Parameter	Nilai
n-gram	(1,2)
max_features	5000
min_df	2
stopword	Indonesian stopwords list

Tabel 7 menunjukkan parameter yang digunakan dalam proses ekstraksi fitur menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Parameter n-gram (1,2) digunakan untuk menangkap kata tunggal (unigram) dan pasangan kata (bigram) yang dapat merepresentasikan konteks linguistik dalam teks. Parameter max\_features membatasi jumlah fitur yang digunakan dalam model, sedangkan min\_df menentukan frekuensi minimum kemunculan kata dalam dokumen agar dapat dipertimbangkan sebagai fitur. Selain itu, digunakan Indonesian stopwords list untuk menghapus kata-kata umum yang tidak memiliki kontribusi signifikan terhadap proses klasifikasi. Pengaturan parameter ini bertujuan untuk menghasilkan representasi numerik teks yang lebih relevan dan efektif dalam proses pemodelan machine learning. TF-IDF digunakan untuk menghitung bobot kata dalam dokumen. Persamaannya adalah:

$$TF\text{-}IDF(t,d)=TF(t,d)\times IDf(t) \tag{1}$$

dengan:

$$TF(t,d) = \frac{f_{t,d}}{\sum f_{t,d}} \tag{2}$$

$$IDF(t) = \log \frac{N}{df_t} \tag{3}$$

Dalam konteks pengolahan teks, simbol (t) merepresentasikan kata atau term, sedangkan (d) menunjukkan dokumen. Notasi (f<sub>t,d</sub>) digunakan untuk menyatakan frekuensi kemunculan suatu kata (t) di dalam dokumen (d).



Selanjutnya, (N) melambangkan jumlah keseluruhan dokumen dalam suatu koleksi atau korpus. Adapun (df) menunjukkan jumlah dokumen yang mengandung kata (t), yang biasanya digunakan untuk mengukur seberapa umum atau spesifik suatu kata dalam keseluruhan dokumen. TF-IDF digunakan untuk merepresentasikan teks ke dalam bentuk vektor numerik dengan memberikan bobot lebih tinggi pada kata yang sering muncul dalam dokumen tetapi jarang muncul dalam keseluruhan korpus.

## 2.6. Pemodelan Random Forest

Model klasifikasi dibangun menggunakan algoritma Random Forest, yaitu metode ensemble learning yang menggabungkan beberapa pohon keputusan untuk meningkatkan akurasi prediksi. Dataset dibagi menjadi data latih dan data uji.

**Tabel 8.** Pembagian Dataset

Dataset	Proporsi
Data latih	80%
Data uji	20%

Tabel 8 menunjukkan pembagian dataset yang digunakan dalam penelitian untuk proses pelatihan dan pengujian model klasifikasi. Dataset dibagi menjadi dua bagian, yaitu data latih sebesar 80% yang digunakan untuk melatih model machine learning, dan data uji sebesar 20% yang digunakan untuk mengevaluasi kinerja model terhadap data yang belum pernah dipelajari sebelumnya. Pembagian ini bertujuan untuk memastikan bahwa model yang dibangun mampu melakukan generalisasi dengan baik dalam mengklasifikasikan data baru.

## 2.7. Optimasi Hyperparameter

Untuk meningkatkan performa model, dilakukan hyperparameter tuning menggunakan Grid Search.

**Tabel 9.** Parameter Tuning Random Forest

Parameter	Nilai yang diuji
n_estimators	100, 200, 300
max_depth	10, 20, None
min_samples_split	2, 5, 10
max_features	sqrt, log2

Tabel 9 menunjukkan parameter yang digunakan dalam proses hyperparameter tuning pada algoritma Random Forest. Proses tuning dilakukan untuk memperoleh kombinasi parameter yang paling optimal dalam meningkatkan performa model klasifikasi. Parameter yang diuji meliputi n\_estimators yang menentukan jumlah pohon keputusan dalam model, max\_depth yang mengatur kedalaman maksimum setiap pohon, min\_samples\_split yang menentukan jumlah minimum sampel untuk melakukan pemisahan node, serta max\_features yang mengatur jumlah fitur yang dipertimbangkan pada setiap proses pemisahan. Pengujian beberapa nilai parameter ini bertujuan untuk menemukan konfigurasi model yang menghasilkan performa klasifikasi terbaik.

## 2.8. Evaluasi Model

Evaluasi model dilakukan menggunakan confusion matrix dan beberapa metrik evaluasi klasifikasi.

**Tabel 10.** Metrik Evaluasi Model

Metrik	Fungsi
Accuracy	Mengukur tingkat ketepatan prediksi
Precision	Mengukur ketepatan prediksi per kelas
Recall	Mengukur kemampuan model menemukan data dalam suatu kelas
F1-Score	Rata-rata harmonis precision dan recall

Tabel 10 menunjukkan metrik evaluasi yang digunakan untuk mengukur kinerja model klasifikasi dalam penelitian ini. Evaluasi model dilakukan menggunakan beberapa metrik umum dalam klasifikasi machine learning, yaitu accuracy, precision, recall, dan F1-score. Accuracy digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan, precision untuk menilai ketepatan model dalam memprediksi suatu kelas tertentu, recall untuk mengukur kemampuan model dalam menemukan seluruh data yang termasuk dalam suatu kelas, sedangkan F1-score merupakan rata-rata harmonis antara precision dan recall yang memberikan gambaran keseimbangan performa model. Penggunaan beberapa metrik ini bertujuan untuk memberikan evaluasi yang lebih komprehensif terhadap kinerja model klasifikasi.

a. Persamaan Accuracy

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

b. Persamaan Precision



$$Precision = \frac{TP}{TP+FP} \quad (5)$$

c. Persamaan Recall

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

d. Persamaan F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

Evaluasi kinerja model klasifikasi dilakukan menggunakan beberapa metrik evaluasi, yaitu accuracy, precision, recall, dan F1-score. Metrik accuracy digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan dengan membandingkan jumlah prediksi yang benar terhadap total data yang diuji. Dalam perhitungannya digunakan empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN).

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Hasil

##### 3.1.1. Karakteristik Dataset

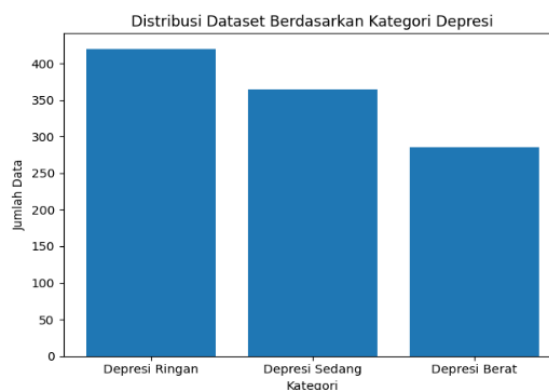
Dataset diperoleh melalui proses scraping pada media sosial X dengan jumlah awal 36.081 unggahan. Setelah dilakukan proses pembersihan data berupa penghapusan duplikasi, penyaringan bahasa, serta eliminasi konten yang tidak relevan, jumlah dataset yang digunakan dalam penelitian ini menjadi 1.070 unggahan teks.

Proses seleksi dataset menghasilkan tingkat penyaringan yang cukup tinggi (lebih dari 97% data awal tidak digunakan). Kondisi ini menunjukkan bahwa sebagian besar unggahan hasil scraping tidak secara langsung merepresentasikan ekspresi emosional yang relevan dengan tujuan penelitian. Dataset akhir cenderung memiliki karakteristik linguistik yang lebih homogen karena hanya menyisakan teks yang berkaitan dengan ekspresi emosional tertentu. Distribusi dataset pada setiap kategori depresi ditunjukkan pada Tabel berikut.

**Tabel 11.** Distribusi Dataset

Kategori	Jumlah Data	Persentase
Depresi ringan	420	39%
Depresi sedang	365	34%
Depresi berat	285	27%
Total	1.070	100%

Distribusi ini menunjukkan bahwa kategori depresi ringan memiliki jumlah data terbesar, sedangkan kategori depresi berat memiliki jumlah data paling sedikit. Perbedaan distribusi ini dapat mempengaruhi proses pembelajaran model karena beberapa kelas memiliki jumlah representasi yang lebih dominan dibandingkan kelas lainnya.



**Gambar 2.** Grafik Distribusi Dataset

Gambar 2 menunjukkan distribusi dataset berdasarkan kategori indikasi depresi. Kategori depresi ringan memiliki jumlah data terbesar, diikuti depresi sedang dan depresi berat. Distribusi ini menunjukkan adanya variasi representasi data pada setiap kategori yang digunakan dalam proses pelatihan dan evaluasi model klasifikasi.

##### 3.1.2. Analisis Ekstraksi Fitur TF-IDF

Representasi numerik teks diperoleh menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) sebagaimana dijelaskan pada bagian metode. Metode ini memberikan bobot pada kata berdasarkan frekuensi kemunculannya dalam dokumen serta tingkat kelangkaannya dalam seluruh korpus teks.



Tabel 12. Contoh Kata dengan Bobot TF-IDF Tinggi

Kategori	Kata Dominan
Depresi ringan	sedih, capek, lelah
Depresi sedang	putus asa, stres, sendiri
Depresi berat	mati, bunuh diri, menyerah

Secara linguistik, kata-kata yang berkaitan dengan ide kematian atau keputusan lebih banyak muncul pada kategori depresi berat, sedangkan kata yang berkaitan dengan kelelahan emosional lebih sering muncul pada kategori ringan dan sedang. Namun demikian, pola ini juga dapat dipengaruhi oleh proses pelabelan manual dataset. Apabila kata tertentu secara konsisten diasosiasikan dengan label tertentu oleh pelabel, maka model akan dengan mudah mempelajari pola tersebut tanpa benar-benar memahami konteks emosional yang lebih kompleks.

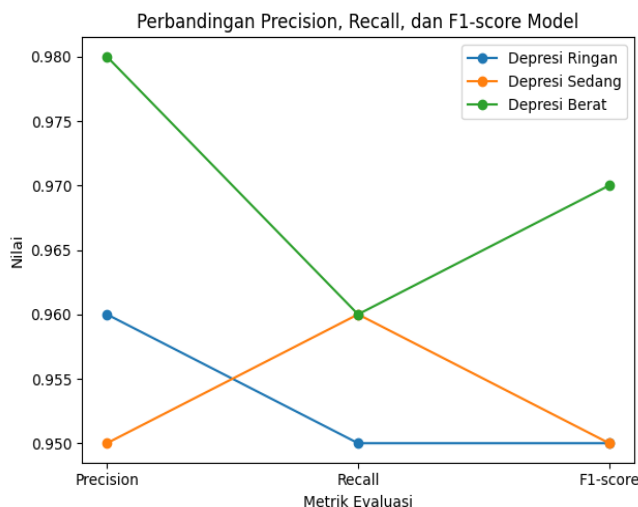
### 3.1.3. Klasifikasi Model Random Forest

Dataset dibagi menjadi 80% data latih dan 20% data uji untuk melatih dan mengevaluasi model Random Forest. Hasil evaluasi model ditunjukkan pada Tabel berikut.

Tabel 13. Hasil Evaluasi Model

Kategori	Precision	Recall	F1-score
Depresi ringan	0.96	0.95	0.95
Depresi sedang	0.95	0.96	0.95
Depresi berat	0.98	0.96	0.97
Accuracy			0.97

Hasil tersebut menunjukkan bahwa model Random Forest mampu mengklasifikasikan teks dengan performa yang tinggi pada seluruh kategori. Namun demikian, angka akurasi sebesar 97% perlu diinterpretasikan secara hati-hati. Dalam klasifikasi teks media sosial yang bersifat tidak terstruktur dan ambigu, nilai akurasi yang sangat tinggi jarang terjadi.



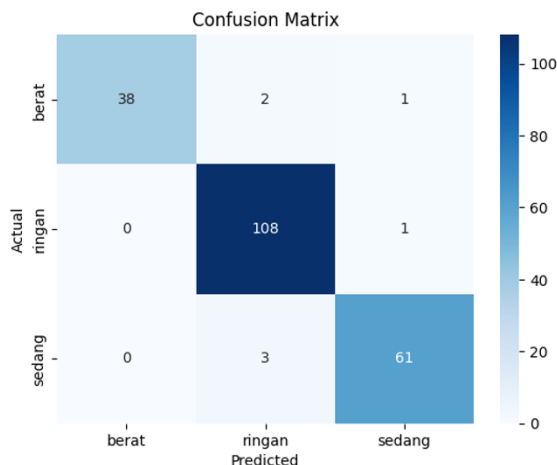
Gambar 3. Grafik Performa Model

Gambar 3 menunjukkan perbandingan nilai precision, recall, dan F1-score pada setiap kategori klasifikasi depresi. Secara umum model Random Forest menunjukkan performa yang tinggi pada seluruh kategori, dengan nilai metrik evaluasi di atas 0,95. Meskipun demikian, performa yang sangat tinggi ini perlu ditafsirkan secara hati-hati karena dapat dipengaruhi oleh karakteristik dataset dan proses pelabelan manual. Confusion matrix digunakan untuk melihat distribusi kesalahan klasifikasi yang dilakukan oleh model.

Tabel 14. Confusion Matrix

Aktual \ Prediksi	Ringan	Sedang	Berat
Ringan	82	3	1
Sedang	2	72	1
Berat	0	2	51

Sebagian besar prediksi model berada pada diagonal utama, yang menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan benar. Kesalahan klasifikasi yang terjadi umumnya berada pada kategori depresi ringan dan depresi sedang, yang secara linguistik memiliki ekspresi emosional yang relatif mirip. Sebaliknya, kategori depresi berat relatif lebih mudah dikenali oleh model karena memiliki karakteristik linguistik yang lebih eksplisit.



**Gambar 4.** Confusion Matrix Hasil Klasifikasi Model Random Forest

Gambar 4 confusion matrix menunjukkan distribusi prediksi model terhadap data aktual pada setiap kategori depresi. Nilai pada diagonal utama menunjukkan jumlah prediksi yang benar pada masing-masing kelas. Model mampu mengklasifikasikan sebagian besar data dengan tepat, terutama pada kategori depresi ringan yang memiliki jumlah prediksi benar paling tinggi. Kesalahan klasifikasi yang terjadi umumnya berada pada kategori depresi sedang dan ringan yang memiliki karakteristik ekspresi linguistik yang relatif mirip. Hasil ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam membedakan tingkat indikasi depresi berdasarkan pola bahasa pada teks media sosial.

### 3.2. Pembahasan

Model klasifikasi berbasis Random Forest mampu mengidentifikasi pola linguistik yang berkaitan dengan ekspresi emosional pada teks media sosial dengan performa yang cukup tinggi. Berdasarkan hasil evaluasi model, nilai accuracy mencapai 97%, dengan nilai precision, recall, dan F1-score pada setiap kategori berada di atas 0,95. Hasil ini menunjukkan bahwa model mampu mempelajari pola linguistik yang membedakan kategori indikasi depresi ringan, sedang, dan berat dalam dataset penelitian. Grafik performa model menunjukkan bahwa nilai precision, recall, dan F1-score relatif konsisten pada setiap kategori. Kategori depresi berat memiliki nilai precision tertinggi, yang menunjukkan bahwa sebagian besar prediksi model pada kategori tersebut merupakan prediksi yang benar. Hal ini dapat terjadi karena kategori depresi berat cenderung memiliki ekspresi linguistik yang lebih eksplisit, seperti penggunaan kata yang berkaitan dengan keputusan atau ide kematian. Sebaliknya, kategori depresi ringan dan sedang memiliki pola bahasa yang lebih bervariasi dan sering kali saling tumpang tindih, sehingga potensi kesalahan klasifikasi pada kedua kategori tersebut relatif lebih tinggi.

Hasil analisis confusion matrix juga menunjukkan bahwa sebagian besar prediksi model berada pada diagonal utama, yang menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan benar. Kesalahan klasifikasi yang terjadi relatif kecil dan umumnya terjadi antara kategori depresi ringan dan depresi sedang. Hal ini menunjukkan bahwa ekspresi emosional yang muncul pada kedua kategori tersebut sering memiliki karakteristik linguistik yang mirip, sehingga model terkadang mengalami kesulitan dalam membedakan keduanya secara tegas. Meskipun hasil evaluasi menunjukkan performa model yang tinggi, nilai akurasi sebesar 97% perlu ditafsirkan secara hati-hati. Dalam konteks klasifikasi teks media sosial yang bersifat informal dan ambigu, nilai performa yang sangat tinggi sering kali dipengaruhi oleh karakteristik dataset yang digunakan. Dalam penelitian ini, proses seleksi dataset dari 36.081 unggahan menjadi 1.070 data menghasilkan korpus teks yang relatif lebih homogen karena sebagian besar data yang tidak relevan telah dieliminasi. Kondisi ini dapat menyebabkan model lebih mudah mengenali pola linguistik tertentu yang telah dominan dalam dataset.

Selain itu, proses pelabelan manual juga dapat mempengaruhi hubungan antara fitur linguistik dan label dataset. Jika pelabel secara konsisten mengasosiasikan kata tertentu dengan kategori tertentu, maka model akan mempelajari pola tersebut secara langsung. Kondisi ini dapat menimbulkan fenomena label bias atau circular reasoning, dimana model tampak memiliki performa tinggi karena mempelajari pola yang sama dengan yang digunakan dalam proses pelabelan. Oleh karena itu, hasil penelitian ini lebih tepat dipahami sebagai upaya eksploratif untuk mengidentifikasi pola bahasa yang berkaitan dengan ekspresi emosional pada media sosial, bukan sebagai sistem diagnosis klinis terhadap depresi. Pendekatan berbasis Natural Language Processing dalam penelitian ini dapat berfungsi sebagai alat analisis awal untuk memahami dinamika ekspresi emosional pada pengguna media sosial, yang pada tahap selanjutnya masih memerlukan validasi lebih lanjut melalui pendekatan psikologis atau klinis.

## 4. KESIMPULAN

Pendekatan analisis teks menggunakan Natural Language Processing (NLP) dan algoritma Random Forest menunjukkan bahwa pola linguistik pada unggahan media sosial dapat dianalisis untuk mengidentifikasi ekspresi emosional yang



muncul dalam teks pengguna. Melalui tahapan pengumpulan data, pembersihan teks, ekstraksi fitur menggunakan TF-IDF, serta proses klasifikasi, model mampu membedakan tiga kategori ekspresi depresi yang digunakan dalam dataset, yaitu depresi ringan, sedang, dan berat. Hasil evaluasi menunjukkan nilai metrik klasifikasi yang tinggi pada data uji, yang menunjukkan bahwa model dapat mempelajari hubungan antara pola bahasa tertentu dan label yang diberikan pada dataset penelitian. Meskipun demikian, hasil tersebut perlu ditafsirkan secara hati-hati. Proses pelabelan data dilakukan berdasarkan interpretasi linguistik terhadap teks media sosial dan tidak merepresentasikan diagnosis klinis terhadap kondisi psikologis individu. Selain itu, indikator Patient Health Questionnaire-9 (PHQ-9) dalam kajian ini digunakan sebagai kerangka konseptual untuk memahami representasi gejala depresi dalam bahasa, bukan sebagai instrumen pengukuran klinis secara langsung. Oleh karena itu, nilai akurasi yang tinggi pada dataset penelitian tidak secara otomatis menunjukkan kemampuan model dalam melakukan deteksi depresi pada konteks klinis nyata. Beberapa keterbatasan juga perlu diperhatikan, antara lain ukuran dataset yang relatif terbatas setelah proses penyaringan data, potensi bias pada proses pelabelan manual, serta belum adanya validasi eksternal terhadap model. Pengembangan selanjutnya disarankan menggunakan dataset yang lebih besar dan beragam serta melibatkan validasi dari ahli kesehatan mental dalam proses pelabelan data sehingga hubungan antara pola bahasa dan kondisi psikologis dapat dianalisis dengan tingkat validitas yang lebih kuat.

## REFERENCES

- [1] A. S. S. Fitriyani and V. R. Setyaning Nastiti, "Pendekatan Linguistik dalam Klasifikasi Emosi Depresi untuk Deteksi Dini Kesehatan Mental di Reddit," *Jurnal Algoritma*, vol. 22, no. 2, Nov. 2025, doi: 10.33364/algoritma/v.22-2.2927.
- [2] O. Christopher Adah and E. Adashona Obiamaka, "An Analysis of the Non-linear Relationship between Test Anxiety and Academic Achievement of Secondary School Students," *International Journal of Depression and Anxiety*, vol. 8, no. 1, Jun. 2025, doi: 10.23937/2643-4059/1710042.
- [3] M. Z. A. Rustam and L. Nurlala, "Gangguan Kecemasan dengan Menggunakan Self Reporting Questionnaire (SRQ-29) di Kota Surabaya," *Jurnal Kesehatan Masyarakat Mulawarman (JKMM)*, vol. 3, no. 1, p. 39, Aug. 2021, doi: 10.30872/jkmm.v3i1.5752.
- [4] M. Y. Dhinora and E. Mailoa, "Analisa Tweet Mahasiswa untuk Deteksi Gejala Depresi dengan Penerapan Natural Language Processing," *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, vol. 6, no. 2, pp. 1193–1211, May 2025, doi: 10.63447/jimik.v6i2.1405.
- [5] F. Apriliani and W. Maharani, "Depression Detection on Social Media Twitter using XLnet Method," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 1, pp. 172–180, Feb. 2023, doi: 10.29100/jupi.v8i1.3345.
- [6] D. Arsyanda, S. Rodiah, and A. S. Rohman, "Peran akun autobase X (twitter) dalam memenuhi kebutuhan informasi followers," *Pustaka Karya: Jurnal Ilmiah Ilmu Perpustakaan dan Informasi*, vol. 13, no. 1, pp. 233–241, Jun. 2025, doi: 10.18592/pk.v13i1.15927.
- [7] D. Arsyanda, S. Rodiah, and A. S. Rohman, "Peran akun autobase X (twitter) dalam memenuhi kebutuhan informasi followers," *Pustaka Karya: Jurnal Ilmiah Ilmu Perpustakaan dan Informasi*, vol. 13, no. 1, pp. 233–241, Jun. 2025, doi: 10.18592/pk.v13i1.15927.
- [8] L. I. Nitami, "Perkembangan Media Sosial Terhadap Perubahan Sosial Masyarakat Di Indonesia Tahun 2000-Sekarang," *KALA MANCA: JURNAL PENDIDIKAN SEJARAH*, vol. 11, no. 2, pp. 69–74, Dec. 2023, doi: 10.69744/kamaca.v11i2.214.
- [9] M. R. Sudrajat and M. Zakariyah, "Penerapan Natural Language Processing dan Machine Learning untuk Prediksi Stres Siswa SMA Berdasarkan Analisis Teks," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, Dec. 2024, doi: 10.47065/bits.v6i3.6180.
- [10] P. Ta, N. Tran, H. Nguyen, and H. D. Nguyen, "Detecting signs of depression on social media: A machine learning analysis and evaluation," *Sustainable Futures*, vol. 10, p. 100827, Dec. 2025, doi: 10.1016/j.sfr.2025.100827.
- [11] V. Oktaviani, N. Rosmawarni, and M. P. Muslim, "Perbandingan Kinerja Random Forest Dan Smote Random Forest Dalam Mendeteksi Dan Mengukur Tingkat Stres Pada Mahasiswa Tingkat Akhir," *Informatik: Jurnal Ilmu Komputer*, vol. 20, no. 1, pp. 43–49, Apr. 2024, doi: 10.52958/iftk.v20i1.9158.
- [12] B. C. Mateus, M. Mendes, J. T. Farinha, and A. Martins, "Hybrid Deep Learning for Predictive Maintenance: LSTM, GRU, CNN, and Dense Models Applied to Transformer Failure Forecasting," *Energies (Basel)*, vol. 18, no. 21, p. 5634, Oct. 2025, doi: 10.3390/en18215634.
- [13] K. Rahayu, V. Fitria, D. Septhya, R. Rahmaddeni, and L. Efrizoni, "Klasifikasi Teks untuk Mendeteksi Depresi dan Kecemasan pada Pengguna Twitter Berbasis Machine Learning," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 108–114, Sep. 2023, doi: 10.57152/malcom.v3i2.780.
- [14] T. T. Widowati and M. Sadikin, "Analisis Sentimen Twitter terhadap Tokoh Publik dengan Algoritma Naive Bayes dan Support Vector Machine," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 11, no. 2, pp. 626–636, Oct. 2021, doi: 10.24176/simet.v11i2.4568.
- [15] N. H. Kim, J. M. Kim, D. M. Park, S. R. Ji, and J. W. Kim, "Analysis of depression in social media texts through the Patient Health Questionnaire-9 and natural language processing," *Digit. Health*, vol. 8, p. 205520762211142, Jan. 2022, doi: 10.1177/20552076221114204.
- [16] F. Apriliani and W. Maharani, "Depression Detection on Social Media Twitter using XLnet Method," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 1, pp. 172–180, Feb. 2023, doi: 10.29100/jupi.v8i1.3345.
- [17] G. F. Situmorang and R. Purba, "Deteksi Potensi Depresi dari Unggahan Media Sosial X Menggunakan IndoBERT," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, pp. 649–661, Sep. 2024, doi: 10.47065/bits.v6i2.5496.
- [18] F. Fatkhurrohman, B. I. Nugroho, and N. Fadillah, "Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 3, pp. 3906–3917, Aug. 2025, doi: 10.31004/riggs.v4i3.2533.



- [19] A. B. Aditya, S. Samsudin, W. P. Rizki, M. Mahendra, and A. Setiawan, “Analisis Sentimen Ulasan Aplikasi Gojek Menggunakan Support Vector Machine Dan Random Forest,” *Jurnal Informatika Terpadu*, vol. 11, no. 2, pp. 134–143, Nov. 2025, doi: 10.54914/jit.v11i2.1884.
- [20] M. Y. Dhinora and E. Mailoa, “Analisa Tweet Mahasiswa untuk Deteksi Gejala Depresi dengan Penerapan Natural Language Processing,” *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 6, no. 2, pp. 1193–1211, May 2025, doi: 10.63447/jimik.v6i2.1405.
- [21] S. P. Aleng, B. J. D. Sitompul, and O. A. Lantang, “Predicting Depression Levels Among Final-Year Students Using CatBoost Algorithm,” *Jurnal Teknik Elektro dan Komputer*, vol. 14, no. 2, pp. 71–82, Dec. 2025, doi: 10.35793/jtek.v14i2.62464.