

Mental Health Classification Using Naïve Bayes and Random Forest Algorithms

Muhammad Jazum Faisti^{1*}, R. Hadapiningradja Kusumodestoni^{2*}, Gentur Wahyu Nyipto Wibowo^{3*}

*Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Nahdlatul Ulama Jepara
jazumfaisty7@gmail.com¹, kusumodestoni@gmail.com², gentur@unisnu.ac.id³

Article Info

Article history:

Received 2025-07-08
Revised 2025-07-25
Accepted 2025-07-30

Keyword:

Mental Health,
Machine Learning,
Naïve Bayes,
Random Forest,
Text Classification.

ABSTRACT

Mental health is a crucial issue affecting individual and societal well-being. This study aims to investigate and compare the performance of Machine Learning algorithms, namely Naïve Bayes and Random Forest, for text-based mental health classification. The dataset used is the Mental Health Corpus from Kaggle, consisting of 27,977 English text messages from online forums, with binary labels (0: no indication of mental disorder, 1: indication of mental disorder) pre-annotated by the dataset creators. Text preprocessing involved lowercasing, negation handling, stopword removal, slang normalization, tokenization, and stemming. Data transformation was performed using TF-IDF. Model evaluation utilized accuracy, precision, recall, and F1-score metrics, along with 5-Fold Cross Validation. Evaluation results indicate high performance for both algorithms. Naïve Bayes achieved 88.7 % accuracy, 84.2 % precision, 95.2 % recall, and 89.3 % F1-score on the test data. Random Forest demonstrated more balanced performance with 89.3 % accuracy, 88.1 % precision, 90.5 % recall, and 89.3 % F1-score. The 5-Fold Cross Validation for Naïve Bayes yielded average scores of 88.8 % accuracy, 84.4 % precision, 94.9 % recall, and 89.3 % F1-score. In contrast, Random Forest showed averages of 89.2 % accuracy, 88.8 % precision, 89.5 % recall, and 89.3 % F1-score. While Naïve Bayes had higher recall, Random Forest exhibited the best overall performance, considering the combination of accuracy, precision, and stable generalization, making it more effective for mental health text classification.



This is an open access article under the [CC-BY-SA](#) license.

I. PENDAHULUAN

Kesehatan mental memainkan peran penting dalam menunjang kesejahteraan baik secara individu maupun sosial. Menurut definisi dari World Health Organization (WHO), kesehatan mental merupakan kondisi sejahtera ketika seseorang mampu menyadari kemampuan dirinya, mengatasi tekanan hidup, bekerja secara produktif, serta berkontribusi kepada lingkungannya. Namun demikian, di tengah kompleksitas kehidupan modern, angka kejadian gangguan kesehatan mental menunjukkan peningkatan yang cukup drastis. [1]. Berdasarkan laporan WHO tahun 2015, lebih dari 300 juta orang di seluruh dunia sekitar 4,4 % populasi global mengalami depresi, dengan kontribusi sebesar 7,5 % terhadap total disabilitas global. Sebagian besar kasus tersebut ditemukan di kawasan Asia Tenggara, termasuk Indonesia,

menjadikan isu ini sebagai masalah kesehatan masyarakat yang krusial dan lintas sektor [2].

Faktor-faktor seperti tekanan akademik, krisis ekonomi, isolasi sosial, serta paparan media digital turut berkontribusi terhadap memburuknya kondisi psikologis individu [3]. Gangguan seperti depresi, kecemasan, dan stres kronis bukan hanya berdampak pada kesehatan individu, tetapi juga menimbulkan konsekuensi sosial dan ekonomi termasuk penurunan produktivitas, isolasi sosial, hingga risiko kehilangan pekerjaan [4]. Sayangnya, masih banyak individu yang tidak mendapatkan penanganan yang memadai akibat keterbatasan akses layanan dan stigma sosial [5].

Dalam menghadapi kompleksitas ini, teknologi informasi dan kecerdasan buatan (Artificial Intelligence/AI) telah menjadi solusi potensial dalam mendukung deteksi dan diagnosis awal gangguan kesehatan mental. Machine

learning, sebagai cabang dari AI, menawarkan pendekatan yang mampu mengidentifikasi pola kompleks dalam data psikologis, khususnya melalui teknik klasifikasi [6]. Berbagai studi sebelumnya telah menerapkan teknik klasifikasi machine learning untuk mendeteksi gangguan kesehatan mental secara dini. Salah satu penelitian menerapkan algoritma Naïve Bayes dan Decision Tree untuk mengklasifikasikan 256 data kuesioner yang dikumpulkan dari responden berusia remaja. Hasilnya, Naïve Bayes menunjukkan performa terbaik dengan akurasi 89,03 % dan nilai kappa 0,856. Temuan ini menegaskan bahwa metode klasifikasi mampu mengenali pola gangguan mental secara efektif, dan berpotensi menjadi alat bantu diagnosis awal [7].

Dalam konteks tersebut, algoritma Naïve Bayes dan Random Forest merupakan dua teknik klasifikasi populer yang sering digunakan dalam kasus serupa [8]. Naïve Bayes, dengan pendekatan probabilistik berdasarkan Teorema Bayes, dikenal karena efisiensinya dalam memproses data kategorikal atau teks [9]. Di sisi lain, Random Forest, sebagai metode ensemble berbasis Decision Tree, unggul dalam mengurangi overfitting dan menangani data non-linear serta high-dimensional [10].

Penelitian yang dilakukan ini bertujuan untuk membandingkan performa kedua algoritma tersebut dalam klasifikasi kondisi kesehatan mental, serta algoritma mana yang menunjukkan performa terbaik menggunakan dataset Mental Health Corpus yang diunduh dari platform Kaggle. Penelitian ini berfokus pada klasifikasi biner, yaitu memetakan apakah seseorang mengalami gangguan kesehatan mental atau tidak.

Beberapa penelitian terdahulu menunjukkan efektivitas Algoritma Naïve Bayes dan Random Forest dalam klasifikasi data kesehatan mental. Salah satu penelitian menerapkan Naïve Bayes pada data survei kesehatan mental dan memperoleh akurasi sebesar 72 %, Hal ini mengindikasikan bahwa algoritma tersebut mampu bekerja secara efektif pada data kategorikal, meskipun masih menghadapi kendala dalam menangani distribusi kelas yang tidak seimbang. [11]. Di sisi lain, penelitian lain menunjukkan bahwa algoritma Random Forest yang dioptimasi dengan validasi silang dan penyetelan parameter mampu mencapai akurasi 90,83 %, recall 90,83 %, dan precision 93,25 % [12].

Pemilihan algoritma Naïve Bayes dan Random Forest dalam penelitian ini didasarkan pada pertimbangan efisiensi, interpretabilitas, dan karakteristik data yang digunakan. Naïve Bayes dipilih karena kemampuannya dalam menangani data teks dengan representasi sederhana seperti TF-IDF dan performanya yang baik pada klasifikasi awal. Sementara itu, Random Forest dipilih sebagai perwakilan algoritma ensemble learning yang mampu menangkap pola kompleks dan memberikan hasil yang stabil.

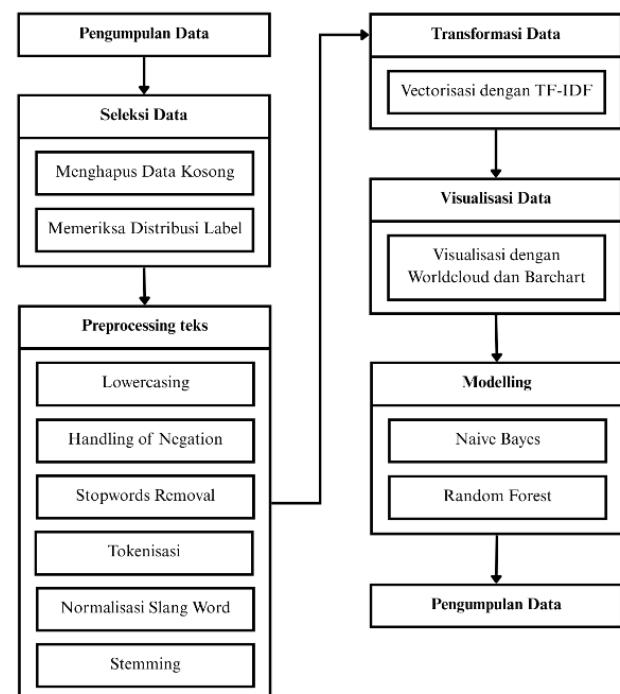
Temuan-temuan ini memperkuat relevansi pemilihan kedua algoritma sebagai objek perbandingan dalam penelitian ini. Meskipun demikian, penelitian-penelitian sebelumnya seringkali berfokus pada data kuesioner atau survei yang terstruktur. Penelitian ini berupaya mengisi kesenjangan

dengan mengeksplorasi efektivitas kedua algoritma pada data teks tidak terstruktur yang berasal dari forum daring, yang merefleksikan ekspresi alami pengguna dan berpotensi menjadi sumber informasi yang kompleks.

Berdasarkan latar belakang tersebut, penelitian ini difokuskan pada evaluasi dan perbandingan performa dua algoritma klasifikasi, yaitu Naïve Bayes dan Random Forest, dalam memproses data terkait kesehatan mental. Evaluasi dilakukan menggunakan sejumlah metrik umum dalam machine learning, seperti akurasi, presisi, recall, dan F1-score. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem skrining berbasis machine learning yang tidak hanya akurat dan efisien, tetapi juga mudah diakses oleh masyarakat umum serta tenaga profesional di bidang kesehatan.

II. METODE

Untuk memudahkan pemahaman terhadap tahapan penelitian yang dilakukan, disajikan diagram alur proses secara sistematis. Diagram ini menggambarkan langkah-langkah mulai dari pengumpulan data hingga evaluasi model klasifikasi.



Gambar. 1 Alur Proses Penelitian

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diambil dari platform sumber data terbuka, Kaggle. Dengan nama Mental Health Corpus (<https://www.kaggle.com/datasets/reihanenamdar/mental-health-corpus>). Dataset tersebut terdiri dari 27.977 entri dalam bentuk pesan singkat berbahasa Inggris yang

dikumpulkan dari forum daring, di mana pengguna secara anonim membagikan ekspresi, keluhan, atau pendapat yang berhubungan dengan kondisi kesehatan mental. Setiap entri memiliki dua atribut utama, yaitu teks sebagai fitur masukan dan label sebagai target klasifikasi. Label pada dataset bersifat biner, dengan nilai 0 menunjukkan bahwa pengguna tidak mengalami indikasi gangguan mental, dan nilai 1 menunjukkan adanya indikasi gangguan mental. Pelabelan ini tidak dilakukan secara self-report, melainkan sudah tersedia dalam dataset sebagai hasil anotasi oleh penyusun dataset di Kaggle.

Pemilihan dataset ini didasarkan pada pertimbangan jumlah data yang memadai, format yang sesuai untuk pemodelan klasifikasi teks, serta ketersediaan label yang relevan untuk pendekatan supervised learning. Karena data diperoleh dari sumber terbuka dan tidak melibatkan partisipasi langsung subjek manusia, maka penelitian ini tidak memerlukan prosedur etis seperti informed consent atau izin eksperimen.

TABEL I
CONTOH DATA MENTAL HEALTH CORPUS

Text	Label
dear american teens question dutch person hear many get baby hit face...	0
nothing look forward life i don't many reasons keep going...	1
music recommendations i'm looking expand playlist...	0
i'm done trying feel better the reason i'm still alive...	1

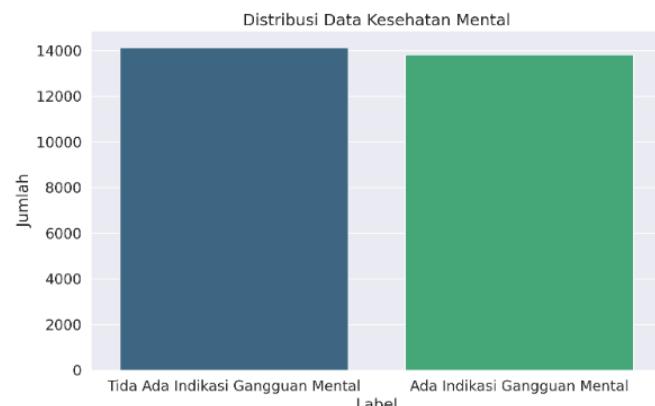
B. Seleksi Data

Setelah proses pengumpulan, data yang diperoleh selanjutnya melalui tahap seleksi awal dalam rangka memastikan bahwa data memiliki kualitas yang memadai dan siap digunakan dalam tahapan pemodelan. Langkah pertama yang dilakukan adalah mengidentifikasi dan menghapus entri yang mengandung nilai kosong (missing values) pada kolom teks maupun label. Hal ini dilakukan untuk menghindari potensi gangguan pada saat pelatihan model yang dapat disebabkan oleh data tidak lengkap.

Setelah tahap pembersihan, dilakukan analisis distribusi label guna mengetahui keseimbangan antara kelas 0 (tidak mengalami gangguan mental) dan kelas 1 (mengalami gangguan mental). Dari hasil analisis pada Gambar 2 menunjukkan bahwa dataset memiliki distribusi yang relatif seimbang, yakni 14.139 entri berlabel 0 dan 13.838 entri berlabel 1, dengan proporsi masing-masing sekitar 51 % dan 49 % dari total 27.977 data. Distribusi ini mengindikasikan bahwa data tidak mengalami ketimpangan kelas (class imbalance) yang signifikan, sehingga evaluasi model menggunakan metrik seperti akurasi tetap relevan dan tidak bias terhadap salah satu kelas.

Namun demikian, untuk memastikan evaluasi model tetap objektif dan tidak bias terhadap satu kelas, penelitian ini tidak hanya menggunakan akurasi sebagai metrik, tetapi juga menambahkan precision, recall, F1-score, serta kurva ROC

dan nilai AUC sebagai indikator performa model secara menyeluruh.



Gambar. 2 Visualisasi Data Kesehatan Mental

C. Teks Preprocessing

Proses pra-pemrosesan teks dilakukan untuk menyarung dan merapikan data mentah sehingga dapat diubah menjadi representasi numerik yang sesuai untuk diproses oleh algoritma machine learning. Proses ini penting untuk mengurangi noise, menghilangkan elemen tidak relevan, dan menstandarkan format teks agar representasi fitur menjadi lebih efektif [13]. Beberapa teknik yang diterapkan dalam penelitian ini dijelaskan sebagai berikut [14]:

- *Lowercasing*: Mengubah seluruh teks menjadi huruf kecil untuk menghindari redundansi, sehingga "Depression" dan "depression" dianggap sama.
- *Handling of Negation*: Menggabungkan kata negasi (seperti "not", "don't") dengan kata setelahnya menjadi satu token (misalnya "not happy" menjadi "not_happy") untuk mempertahankan makna negatif.
- *Stopword Removal*: Menghapus kata-kata umum yang tidak penting, seperti "and", "the", dan "is", menggunakan daftar stopword dari NLTK untuk mengurangi fitur tidak informatif.
- *Slang Normalization*: Menormalkan kata tidak baku atau singkatan (misalnya "im" menjadi "I am") ke bentuk standar untuk menjaga makna.
- *Tokenisasi*: Memecah kalimat menjadi unit kata (token) untuk analisis linguistik.
- *Stemming*: Mengembalikan kata ke bentuk dasarnya menggunakan Porter Stemmer, misalnya "running", "runs", dan "ran" menjadi "run".

TABEL II
CONTOH TAHAPAN TEXT PREPROCESSING

No	Tahapan	Contoh Sebelum	Contoh Sesudah
1	Lowercasing	"I Feel SAD Today"	"i feel sad today"
2	Handling of Negation	"i do not feel happy"	"i not_feel happy"

3	Stopword Removal	"i am very sad today"	"sad today"
4	Slang Normalization	"idk what to do"	"i do not know what to do"
5	Tokenization	"i feel sad"	["i", "feel", "sad"]
6	Stemming	"feeling", "feels", "felt"	"feel"

D. Transformasi Data

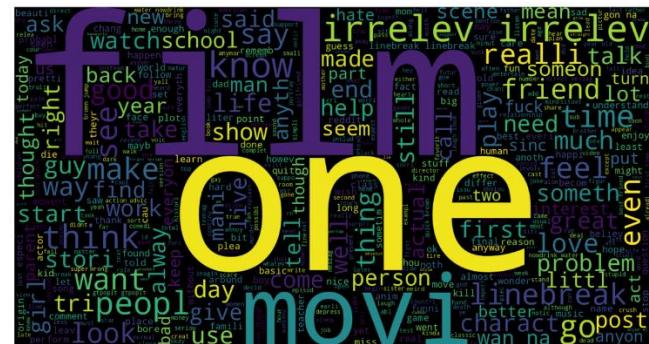
Transformasi data merupakan proses mengubah teks yang telah melalui tahap pra-pemrosesan menjadi format numerik, sehingga dapat digunakan oleh algoritma machine learning. Proses ini sangat penting dalam bidang pemrosesan bahasa alami (Natural Language Processing/NLP), teks tidak dapat langsung digunakan sebagai input ke dalam model klasifikasi. Oleh karena itu, diperlukan metode untuk mengubah data tidak terstruktur (teks) menjadi vektor numerik berdimensi tetap yang mencerminkan bobot pentingnya kata dalam setiap dokumen [15].

Dalam penelitian ini, pendekatan Term Frequency-Inverse Document Frequency (TF-IDF) digunakan sebagai metode transformasi data. Teknik ini memberikan bobot pada setiap kata berdasarkan frekuensinya dalam dokumen, sambil mempertimbangkan kelangkaan kemunculannya di seluruh koleksi dokumen [16]. Keunggulan utama TF-IDF adalah kemampuannya mengurangi pengaruh kata-kata umum seperti "dan", "adalah", atau "yang", yang dianggap tidak memiliki nilai informasi spesifik karena frekuensinya yang tinggi. Melalui komponen Inverse Document Frequency, bobot kata-kata tersebut ditekan menjadi kecil, sehingga model dapat menekankan kata-kata yang membedakan satu dokumen dari dokumen lainnya [17]. Hal ini menjadikan TF-IDF sangat efektif dalam menangkap fitur penting dari teks untuk keperluan klasifikasi atau analisis lanjutan.

Sebelum transformasi, data dibagi menggunakan teknik train-test split (80 % data latih, 20 % data uji). TF-IDF kemudian diterapkan dengan batas maksimal 2.500 fitur dan mengabaikan kata yang muncul kurang dari dua kali untuk mengurangi sparsity. Hasilnya adalah matriks fitur berdimensi tetap yang merepresentasikan setiap teks sebagai vektor dengan bobot TF-IDF yang relevan.

E. Visualisasi Data

E. *Visualisasi Data*
Visualisasi data adalah langkah untuk mengubah data mentah menjadi representasi grafis guna mempermudah analisis dan pemahaman terhadap pola, struktur, serta tren yang terkandung dalam data [18]. Dalam penelitian ini, visualisasi digunakan untuk menggambarkan distribusi label antara individu yang terindikasi mengalami gangguan kesehatan mental dan yang tidak, serta untuk mengidentifikasi karakteristik teks seperti panjang kalimat dan frekuensi kata yang dominan. Representasi visual seperti bar chart dan wordcloud dipilih karena kemampuannya menyajikan informasi secara ringkas dan komunikatif.



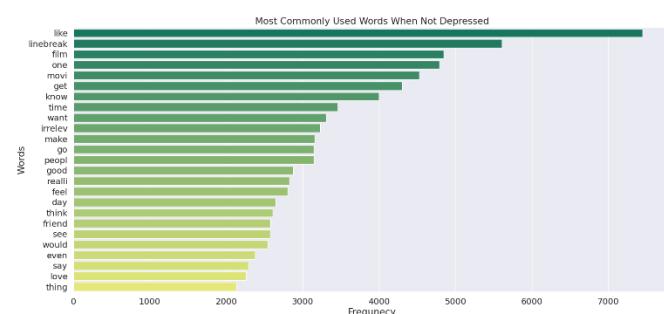
Gambar. 3 Wordcloud Non Gangguan Mental

Gambar 3 menunjukkan frekuensi kata-kata yang paling sering muncul dalam teks-teks yang diklasifikasikan sebagai "Tidak Ada Indikasi Gangguan Mental" (Label 0). Bawa ukuran kata menunjukkan seberapa sering kata tersebut muncul, dengan kata-kata seperti "movie", "friend", "film", "story", dan "one" mendominasi, merefleksikan tema-tema hiburan, interaksi sosial normal, dan rutinitas positif.



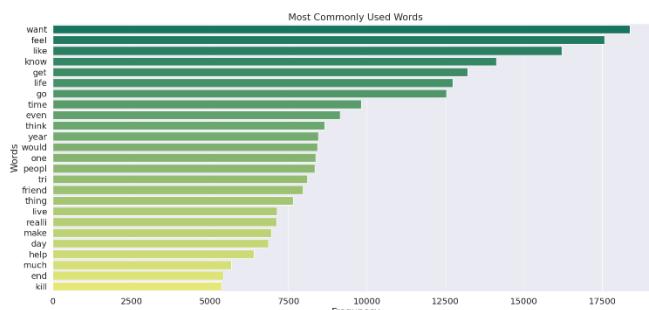
Gambar. 4 Wordcloud Gangguan Mental

Pada Gambar 4 menampilkan kata-kata yang paling dominan dalam teks-teks yang diklasifikasikan sebagai "Ada Indikasi Gangguan Mental" (Label 1). Kata-kata seperti "know", "want", "feel", "life", "problem", dan "think" berukuran besar, mengindikasikan fokus pada pengalaman pribadi yang emosional, keinginan untuk memahami, serta isu-isu eksistensial atau sosial yang sering dikaitkan dengan kondisi kesehatan mental.



Gambar. 5 Kata Umum Non Gangguan Mental

Gambar 5 menunjukkan distribusi frekuensi kata-kata umum pada pengguna yang tidak terindikasi gangguan mental (Label 0), di mana kata-kata seperti "like", "film", dan "movie" memiliki frekuensi tinggi, mengindikasikan diskusi yang berpusat pada aktivitas sehari-hari, hiburan, atau opini umum yang bersifat netral atau positif.



Gambar. 6 Frekuensi Kata Tertinggi (Umum)

Gambar 6 menyajikan kata-kata dengan frekuensi kemunculan tertinggi di seluruh dataset, seperti "want", "feel", dan "like", menegaskan bahwa ekspresi diri, keinginan, dan perasaan merupakan komponen sentral dalam komunikasi pengguna terkait kondisi mental mereka.

F. K-Fold Cross Validation

K-Fold Cross Validation merupakan teknik validasi silang yang bertujuan untuk mengevaluasi keandalan dan kemampuan generalisasi model machine learning terhadap data yang belum pernah dilihat sebelumnya. Metode ini membagi dataset menjadi k bagian yang ukurannya setara. Model kemudian dilatih pada k-1 bagian dan diuji pada bagian yang tersisa. Proses ini dilakukan sebanyak k kali dengan pergantian bagian uji, dan hasil dari seluruh iterasi dirata-ratakan guna memperoleh metrik evaluasi yang lebih konsisten. [19].

Pada penelitian ini, validasi model dilakukan menggunakan skema K-Fold Cross Validation dengan jumlah fold ditetapkan sebanyak lima ($k = 5$), yang memastikan distribusi label seimbang pada setiap fold. Nilai $k=5$ dipilih karena merupakan konfigurasi umum yang memberikan estimasi performa model yang stabil tanpa beban komputasi yang terlalu tinggi, sekaligus memastikan setiap fold memiliki ukuran yang cukup representatif. Selain itu, pendekatan ini dipilih untuk menghindari overfitting dan memberikan estimasi performa model yang lebih akurat, terutama mengingat data bersifat tidak berurutan dan memiliki proporsi kelas yang hampir seimbang.

G. Algoritma Naïve Bayes

Metode Naïve Bayes merupakan pendekatan klasifikasi yang mengacu pada prinsip Teorema Bayes, yang dikembangkan oleh ilmuwan Inggris bernama Thomas Bayes. Teorema tersebut memberikan dasar matematis untuk mengestimasi kemungkinan suatu peristiwa berdasarkan pengetahuan terhadap peristiwa sebelumnya [20]. Teorema

ini memperkirakan probabilitas suatu peristiwa dengan mempertimbangkan informasi yang diperoleh dari kondisi lingkungan serta data sampel yang tersedia. Dalam penerapannya, Teorema Bayes memanfaatkan informasi awal (prior knowledge) sebagai landasan untuk menentukan variabel-variabel yang relevan dalam proses estimasi parameter baru. [21]. Secara matematis, Teorema Bayes dirumuskan sebagai berikut [22]:

$$P(H|X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} \quad (1)$$

Keterangan :

X = Data yang kelasnya belum diketahui sebelumnya,
 C_i = Hipotesis data X termasuk ke dalam kelas tertentu,
 $P(C_i|X)$ = Probabilitas kelas C terhadap fitur X (posterior),
 $P(C_i)$ = Probabilitas awal dari hipotesis C_i ,
 $P(X|C_i)$ = Probabilitas kemunculan fitur X dalam kelas C,
 $P(X)$ = Probabilitas X.

Pada penelitian ini, digunakan varian Multinomial Naïve Bayes yang merupakan jenis yang paling umum digunakan untuk klasifikasi teks. Algoritma ini sangat sesuai untuk data berbasis frekuensi kata dan mampu menangani vektor fitur hasil transformasi TF-IDF secara efektif. Kelebihan utama dari algoritma ini adalah kecepatannya dalam pelatihan model dan ketahanannya terhadap noise pada data yang besar.

H. Algoritma Random Forest

Random Forest adalah metode pembelajaran mesin berbasis ensemble yang menggabungkan sejumlah decision tree yang dibangun secara acak. Tujuan utamanya adalah untuk meningkatkan performa prediksi sekaligus meminimalkan overfitting yang sering menjadi kelemahan pada model decision tree tunggal. Dalam penerapannya, Random Forest menyatukan hasil dari semua pohon dengan menggunakan mekanisme voting mayoritas untuk tugas klasifikasi atau menghitung nilai rata-rata untuk tugas regresi [23].

Setiap pohon dalam Random Forest dilatih menggunakan teknik bootstrap aggregating (bagging), yaitu pengambilan subset data latih secara acak dengan pengembalian. Pada setiap percabangan pohon (split), hanya sebagian fitur yang dipilih secara acak untuk dipertimbangkan, menghasilkan model yang beragam namun saling melengkapi sehingga prediksi lebih stabil dan akurat. Pemilihan split terbaik didasarkan pada information gain, yaitu pengurangan entropy setelah pemisahan. Fitur dengan information gain tertinggi dipilih karena paling efektif memisahkan data berdasarkan kelas [24]. Berikut rumus pemisahan simpul didasarkan pada pencarian entropy [10]:

$$\text{Entropy}(S) = \sum_{i=1}^n -\log_2(P_i) \quad (2)$$

Keterangan :

S = Himpunan kasus,

n = Jumlah partisi S,
 p_i = Proporsi dari S_i terhadap S.

Pemilihan algoritma Random Forest dalam studi ini didasarkan pada kemampuannya yang unggul dalam mengolah data dengan dimensi yang kompleks dan jumlah fitur yang tinggi, dan non-linear seperti teks hasil ekstraksi fitur TF-IDF. Selain itu, algoritma ini cenderung lebih tahan terhadap outlier dan noise, serta dapat menangani data dengan interaksi kompleks antar fitur tanpa memerlukan banyak parameter penyetelan. Random Forest juga menyediakan estimasi pentingnya fitur (feature importance), yang dapat digunakan untuk memahami kontribusi kata tertentu terhadap proses klasifikasi.

I. Evaluasi

Evaluasi model bertujuan untuk menilai tingkat kemampuan algoritma klasifikasi dalam menghasilkan prediksi yang akurat. Dalam penelitian ini, metode utama yang digunakan untuk menilai performa model adalah confusion matrix, yang menggambarkan perbandingan antara hasil prediksi model dan label sebenarnya pada data uji. Confusion matrix ini memberikan gambaran mengenai jumlah prediksi yang sesuai dengan kebenaran serta kesalahan klasifikasi yang dilakukan oleh model [25].

TABEL III
KONSEP CONFUSION MATRIX

		True Class	
Predicted Class	Class 1	Class 1	Class 2
	Class 1	True Positive	False Negative
	Class 2	False Positive	True Negative

Keterangan :

True Positive (TP) = Data positif yang diprediksi positif.
 True Negative (TN) = Data negatif yang diprediksi negatif.
 False Positive (FP) = Data negatif yang diprediksi positif.
 False Negative (FN) = Data positif yang diprediksi negatif.

Berdasarkan susunan confusion matrix, sejumlah metrik evaluasi digunakan untuk memberikan penilaian yang lebih mendalam terhadap performa model. Metrik-metrik yang dimanfaatkan dalam penelitian ini dijabarkan sebagai berikut[26]:

- Akurasi (Accuracy) merupakan metrik yang digunakan untuk menghitung proporsi prediksi yang tepat dibandingkan dengan total jumlah data. Rumus perhitungannya disajikan dalam persamaan berikut 3.

$$Accuracy = \frac{(TP + TN)}{TP + FP + FN + TN} \quad (3)$$

- Presisi (Precision) mengukur seberapa tepat prediksi positif yang dilakukan oleh model, yaitu seberapa banyak prediksi positif yang benar-benar relevan. Dirumuskan pada persamaan 4.

$$Precision = \frac{(TP)}{TP + FP} \quad (4)$$

- Recall digunakan untuk menilai sejauh mana model mampu mengenali seluruh data yang termasuk dalam kategori positif, yaitu dengan menghitung proporsi kasus positif yang berhasil terdeteksi secara benar oleh model. Dirumuskan pada persamaan 5.

$$Recall = \frac{(TP)}{TP + FN} \quad (5)$$

- F1-Score adalah nilai rata-rata harmonik dari precision dan recall, yang digunakan untuk memberikan evaluasi yang seimbang antara kedua metrik tersebut, terutama dalam kondisi distribusi data yang tidak seimbang. Dirumuskan pada persamaan 6.

$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (6)$$

Dengan menggunakan metrik-metrik di atas, performa dari algoritma Naïve Bayes dan Random Forest dapat dianalisis dan dibandingkan secara kuantitatif. Evaluasi ini tidak hanya mengukur ketepatan prediksi, tetapi juga mempertimbangkan konsekuensi dari kesalahan klasifikasi, yang sangat penting dalam konteks deteksi kondisi kesehatan mental.

III. HASIL DAN PEMBAHASAN

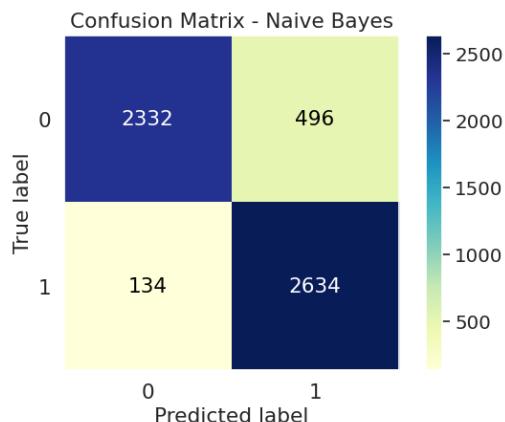
Pada tahap ini menyajikan hasil evaluasi terhadap dua model klasifikasi kesehatan mental yang telah dikembangkan, yaitu Naïve Bayes dan Random Forest. Evaluasi dilakukan untuk menilai sejauh mana model mampu mengklasifikasikan status mental pengguna berdasarkan data teks yang telah melalui proses preprocessing dan transformasi. Analisis performa model dilakukan dengan membandingkan metrik evaluasi standar serta visualisasi untuk mendukung interpretasi hasil.

A. Hasil Evaluasi Model Naïve Bayes

Model Naïve Bayes menunjukkan performa yang sangat baik dalam mengklasifikasikan data kesehatan mental berbasis teks. Evaluasi dilakukan menggunakan data uji dan menghasilkan nilai metrik seperti ditampilkan pada Tabel 4.

TABEL IV
HASIL EVALUASI MODEL NAÏVE BAYES

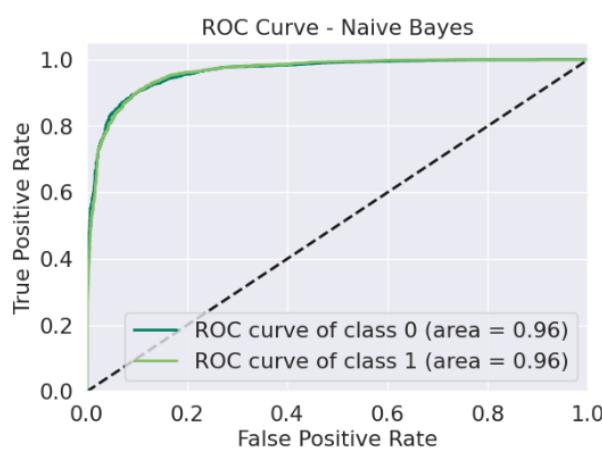
Metrik	Nilai
Akurasi	0.887
Presisi	0.842
Recall	0.952
F1-score	0.893



Gambar. 7 Confusion Matrix Model Naïve Bayes

Berdasarkan confusion matrix pada Gambar 7 dan hasil evaluasi, model Naïve Bayes menunjukkan performa yang baik dengan nilai akurasi sebesar 0,887, precision 0,842, recall 0,952, dan F1-score 0,893. Model berhasil mengklasifikasikan 2.332 data negatif dan 2.634 data positif secara benar, meskipun masih terdapat 496 false positive dan 134 false negative. Tingginya recall menunjukkan kemampuan model dalam mengenali data berindikasi gangguan mental, namun precision yang lebih rendah mengindikasikan kecenderungan model melakukan prediksi positif berlebih. Hal ini berkaitan dengan karakteristik Naïve Bayes yang mengasumsikan independensi fitur, sehingga kata-kata tertentu yang sering muncul dalam kelas gangguan mental dapat memengaruhi prediksi meski konteksnya berbeda.

Selain itu, Gambar 8 yang menampilkan kurva ROC memperlihatkan bahwa model memiliki tingkat diskriminasi yang tinggi, dibuktikan dengan nilai AUC sebesar 0,96 untuk kedua kategori. Nilai ini mencerminkan kemampuan model dalam memisahkan antara kelas 0 dan kelas 1 secara konsisten pada berbagai threshold klasifikasi.



Gambar. 8 ROC Curve Model Naïve Bayes

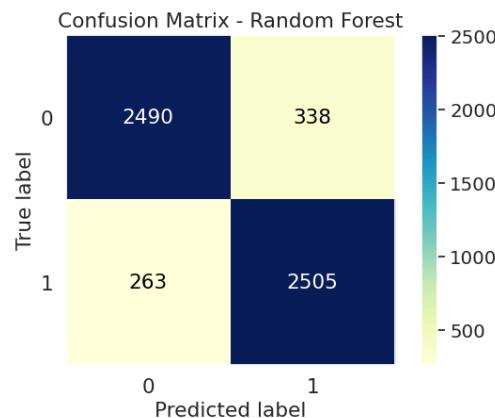
Secara keseluruhan, algoritma Naïve Bayes menunjukkan performa yang solid dengan keseimbangan antara recall dan precision yang baik. Meskipun ada kelemahan pada ketepatan klasifikasi negatif, model ini tetap relevan untuk aplikasi deteksi awal, khususnya ketika prioritas diletakkan pada mengurangi jumlah kasus positif yang terlewat (false negative).

B. Hasil Evaluasi Model Random Forest

Model Random Forest memberikan hasil klasifikasi yang sangat baik pada data uji. Evaluasi menunjukkan bahwa model ini mencapai nilai akurasi, presisi, recall, dan F1-score yang relatif seimbang dan konsisten. sebagaimana ditampilkan pada Tabel 5.

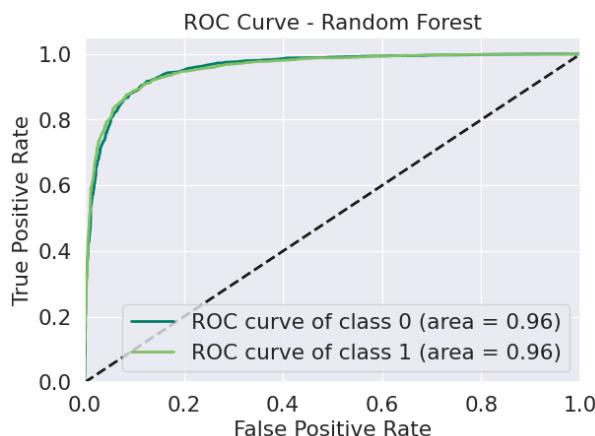
TABEL V
HASIL EVALUASI MODEL RANDOM FOREST

Metrik	Nilai
Akurasi	0,893
Presisi	0,881
Recall	0,905
F1-score	0,893



Gambar. 9 Confusion Matrix Model Random Forest

Berdasarkan confusion matrix pada Gambar 9, model Random Forest mampu mengklasifikasikan 2.490 data negatif dan 2.505 data positif secara tepat, dengan kesalahan klasifikasi berupa 338 false positive dan 263 false negative. Evaluasi metrik menunjukkan performa yang cukup seimbang, dengan precision sebesar 0,881, recall 0,905, F1-score 0,893, dan akurasi keseluruhan mencapai 0,893. Meskipun terdapat sejumlah kesalahan prediksi, jumlahnya relatif kecil dan secara keseluruhan model tetap menunjukkan performa klasifikasi yang stabil.



Gambar. 10 ROC Curve Model Random Forest

Selain itu, performa diskriminatif model terhadap dua kelas ditunjukkan melalui ROC Curve pada Gambar 10, dengan nilai AUC sebesar 0,96 untuk masing-masing kelas. Hal ini mengindikasikan bahwa model mampu membedakan dengan baik antara dua label klasifikasi pada berbagai nilai ambang prediksi.

Secara keseluruhan, model Random Forest menunjukkan kinerja yang handal dalam mendeteksi ekspresi teks terkait kondisi kesehatan mental dan layak digunakan sebagai pendekatan klasifikasi berbasis data teks. Keseimbangan performa Random Forest, terutama pada metrik precision dan recall, menunjukkan kemampuannya dalam menangani kompleksitas data teks dan mengurangi risiko overfitting melalui mekanisme ensemble, menjadikannya pilihan yang robust untuk tugas klasifikasi ini.

C. Validasi Kinerja Model

Validasi model menjadi tahapan krusial dalam menilai konsistensi kinerja algoritma terhadap berbagai variasi data. Pada penelitian ini, digunakan metode K-Fold Cross Validation dengan nilai k sebesar lima. Pendekatan ini memastikan bahwa setiap subset data memiliki distribusi label yang merata, sehingga proses evaluasi model dapat dilakukan secara lebih objektif dan tidak bias terhadap satu skema pembagian data tertentu.

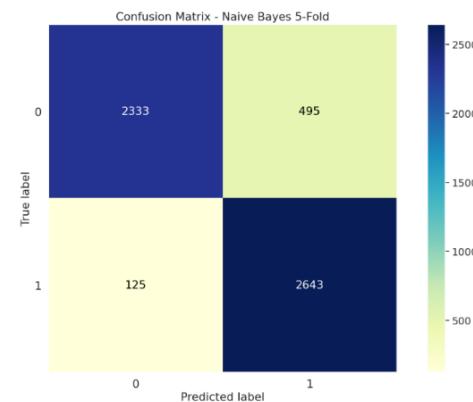
1) Validasi Model Naïve Bayes

Hasil validasi model Naïve Bayes dengan 5-Fold menunjukkan performa yang kuat dan stabil dalam klasifikasi data. Tabel 6 berikut menyajikan rata-rata hasil metrik evaluasi.

TABEL VI
HASIL VALIDASI NAÏVE BAYES (5-FOLD CROSS VALIDATION)

Metrik Evaluasi	Nilai Rata-rata
Accuracy	0.888
Precision	0.844
Recall	0.949
F1-Score	0.893

Hasil validasi model Naïve Bayes dengan skema 5-Fold Cross Validation menunjukkan performa yang konsisten dan andal dalam klasifikasi data teks kesehatan mental. Rata-rata metrik evaluasi yang diperoleh mencakup akurasi sebesar 0,887, precision 0,843, recall 0,949, dan F1-score 0,893. Tingginya nilai recall memperkuat temuan bahwa model sangat efektif dalam mengenali data berlabel positif (indikasi gangguan mental), meskipun precision yang lebih rendah menunjukkan masih adanya false positive yang cukup signifikan. Nilai F1-score yang tinggi menandakan keseimbangan performa yang baik.



Gambar. 11 Confusion Matrix Model Naïve Bayes (5 Fold)

Pada Gambar 11, confusion matrix menunjukkan bahwa model berhasil mengklasifikasikan 2.643 data positif dan 2.333 data negatif secara akurat, dengan kesalahan berupa 495 false positive dan 125 false negative. Kesalahan klasifikasi cenderung terjadi saat model menghadapi teks dengan kata-kata yang sering muncul pada kelas gangguan, tetapi dalam konteks yang sebenarnya netral.

Distribusi prediksi yang dominan benar pada kedua kelas menunjukkan bahwa model Naïve Bayes bekerja cukup andal dalam klasifikasi data teks terkait kesehatan mental, meskipun masih terdapat kecenderungan meningkatkan false positive, yang dapat berimplikasi pada sensitivitas model terhadap ekspresi netral.

2) Validasi Model Random Forest

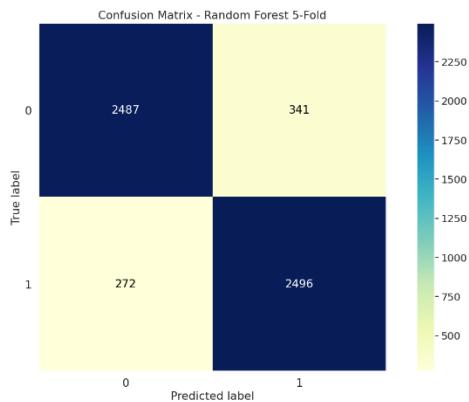
Model Random Forest juga divalidasi menggunakan teknik 5-Fold Cross Validation. Hasil evaluasi selama proses validasi disajikan dalam Tabel 7 berikut:

TABEL VII
HASIL VALIDASI RANDOM FOREST (5-FOLD CROSS VALIDATION)

Metrik Evaluasi	Nilai Rata-rata
Accuracy	0.892
Precision	0.888
Recall	0.895
F1-Score	0.893

Hasil validasi menggunakan 5-Fold Cross Validation menunjukkan bahwa model Random Forest memiliki

performa klasifikasi yang kuat dan konsisten. Rata-rata nilai metrik evaluasi mencakup akurasi sebesar 0,893, precision 0,888, recall 0,898, dan F1-score 0,893. Nilai-nilai ini mencerminkan bahwa model tidak hanya mampu mengenali data dengan indikasi gangguan mental secara akurat, tetapi juga menjaga tingkat kesalahan klasifikasi pada level yang rendah. Keseimbangan antara precision dan recall memperlihatkan bahwa model cukup sensitif dalam mendeteksi kelas positif, tanpa mengorbankan akurasi terhadap kelas negatif.



Gambar. 12 Confusion Matrix Model Random Forest (5 Fold)

Gambar 12 menunjukkan confusion matrix model Random Forest dengan 5-Fold Cross Validation. Model secara tepat mengklasifikasikan 2.487 data negatif dan 2.496 data positif, dengan kesalahan sebesar 341 False Positive dan 272 False Negative. Hasil ini mencerminkan performa yang seimbang dan akurat dalam mendeteksi kedua kelas tanpa bias signifikan. Namun demikian, beberapa kesalahan klasifikasi masih terjadi, terutama pada teks yang menggunakan bahasa implisit, metaforis, atau mengandung nuansa emosional tidak langsung.

Dengan keseluruhan hasil tersebut, dapat disimpulkan bahwa Random Forest merupakan model yang andal dan stabil, baik dalam hal akurasi keseluruhan, sensitivitas terhadap kelas positif, maupun kemampuan membedakan antar kelas secara konsisten dalam skenario validasi berulang.

D. Analisis dan Interpretasi Performa Algoritma

Analisis performa dilakukan dengan membandingkan dua algoritma klasifikasi Naïve Bayes dan Random Forest berdasarkan hasil evaluasi terhadap data uji dan validasi K-Fold. Tujuan dari analisis ini adalah untuk mengevaluasi kelebihan, kekurangan, serta tingkat kecocokan masing-masing model dalam menangani klasifikasi teks yang berkaitan dengan isu kesehatan mental.

1) Perbandingan Ketepatan Umum (Akurasi)

Akurasi memberikan gambaran awal mengenai proporsi prediksi yang benar dari seluruh data uji. Seperti pada Tabel 8, model Random Forest mencatatkan rata-rata akurasi 0,893, sedikit lebih tinggi dari Naïve Bayes yang berada di 0,887.

Meskipun selisihnya kecil, hal ini menunjukkan bahwa Random Forest sedikit lebih konsisten secara umum.

TABEL VIII
PERBANDINGAN AKURASI

Model	Akurasi Data Uji	Akurasi K-Fold
Naïve Bayes	0,887	0,888
Random Forest	0,893	0,892

2) Ketepatan Prediksi Positif (Precision)

Tabel 9 menunjukkan bahwa Random Forest memperoleh nilai precision tertinggi, yaitu 0,888, baik pada pengujian maupun saat validasi. Hal ini menandakan bahwa model memiliki kemampuan yang baik dalam meminimalkan kesalahan klasifikasi data negatif sebagai false positive.

TABEL IX
PERBANDINGAN PRECISION

Model	Precision Uji	Precision K-Fold
Naïve Bayes	0,842	0,844
Random Forest	0,881	0,888

3) Sensitivitas terhadap Kelas Positif (Recall)

Berdasarkan Tabel 10, Naïve Bayes memiliki keunggulan yang jelas dengan recall mencapai 0,949 pada validasi 5-fold, menunjukkan sensitivitas yang tinggi dalam mengidentifikasi kasus dengan indikasi gangguan mental.

TABEL X
PERBANDINGAN RECALL

Model	Recall Uji	Recall K-Fold
Naïve Bayes	0,952	0,949
Random Forest	0,905	0,895

4) Keseimbangan Performa (F1-Score)

Pada Tabel 11, kedua model menunjukkan nilai F1-score yang sama tinggi, yakni 0,893 pada validasi. Meskipun pendekatannya berbeda (Naïve Bayes unggul di recall, Random Forest unggul di precision), keduanya menghasilkan keseimbangan prediksi yang sama kuat.

TABEL XI
PERBANDINGAN F1 - SCORE

Model	F1 Uji	F1 K-Fold
Naïve Bayes	0,893	0,893
Random Forest	0,893	0,893

Meskipun kedua model menunjukkan performa metrik yang tinggi dan relatif seimbang, keduanya masih menghadapi tantangan dalam mengklasifikasikan teks yang bersifat ambigu, implisit, atau bersifat emosional halus. Model Naïve Bayes cenderung menghasilkan false positive karena terlalu bergantung pada kemunculan kata-kata tertentu yang sering muncul di kelas positif, meskipun konteksnya netral atau sarkastik. Di sisi lain, Random Forest lebih konservatif, namun tetap mengalami kesulitan dalam memahami kalimat dengan makna tersirat, metafora, atau narasi panjang yang mengandung emosi kompleks. Kedua

model juga belum mampu menangkap konteks antar kalimat atau dinamika psikologis yang tidak eksplisit, seperti perasaan putus asa yang disampaikan secara tidak langsung. Hal ini menunjukkan bahwa meskipun pendekatan machine learning tradisional dapat memberikan hasil yang kuat secara numerik, mereka tetap memiliki keterbatasan dalam memahami konteks semantik yang lebih dalam, yang justru sering muncul dalam ekspresi individu dengan kondisi kesehatan mental tertentu.

IV. KESIMPULAN

Berdasarkan hasil analisis dan evaluasi kinerja terhadap dua algoritma klasifikasi Naïve Bayes dan Random Forest dapat disimpulkan bahwa keduanya mampu melakukan klasifikasi data teks terkait kesehatan mental dengan akurasi yang tinggi dan stabil. Pada tahap evaluasi menggunakan data uji, algoritma Naïve Bayes berhasil mencapai akurasi sebesar 88,7 %, precision sebesar 84,2 %, recall sebesar 95,2 %, dan F1-score sebesar 89,3 %. Sementara itu, algoritma Random Forest menunjukkan kinerja yang lebih seimbang dengan akurasi sebesar 89,3 %, precision sebesar 88,1 %, recall sebesar 90,5 %, dan F1-score sebesar 89,3 %.

Validasi lebih lanjut menggunakan skema 5-Fold Cross Validation memperkuat temuan tersebut, di mana Naïve Bayes memperoleh rata-rata akurasi sebesar 88,8 %, precision 84,4 %, recall 94,9 %, dan F1-score 89,3 %. Adapun Random Forest mencatatkan rata-rata akurasi sebesar 89,2 %, precision 88,8 %, recall 89,5 %, dan F1-score 89,3 %.

Berdasarkan keseluruhan metrik tersebut, dapat disimpulkan bahwa algoritma Random Forest memiliki performa paling optimal, ditinjau dari kombinasi akurasi, precision, dan keseimbangan antara recall serta kemampuan generalisasi yang stabil pada skema validasi. Meskipun Naïve Bayes menunjukkan keunggulan pada aspek recall yang lebih tinggi, Random Forest menawarkan performa yang lebih proporsional antara sensitivitas dan ketepatan prediksi, sehingga lebih efektif untuk aplikasi klasifikasi kesehatan mental berbasis teks.

Namun demikian, kedua algoritma masih menghadapi beberapa keterbatasan, terutama dalam mengklasifikasikan teks dengan karakteristik bahasa yang ambigu, ekspresi metaforis, atau emosi yang tersirat. Naïve Bayes cenderung menghasilkan false positive akibat ketergantungan pada frekuensi kata tanpa mempertimbangkan konteks, sedangkan Random Forest, meskipun lebih konservatif, masih kesulitan dalam memahami makna semantik yang kompleks dan konteks interkalimat.

Untuk mengatasi keterbatasan tersebut, disarankan penelitian selanjutnya mengeksplorasi pendekatan berbasis deep learning, khususnya model dengan arsitektur Transformer seperti BERT atau RoBERTa yang telah terbukti efektif dalam memahami konteks dan relasi antar kata dalam teks panjang. Peningkatan pada tahap preprocessing dan pengayaan fitur juga menjadi krusial dalam menangani

bahasa implisit dan ekspresi metaforis yang sering muncul dalam diskursus kesehatan mental.

Selain itu, penting untuk dicatat bahwa deteksi kesehatan mental melalui analisis teks memiliki tingkat sensitivitas yang tinggi dan berpotensi menimbulkan risiko apabila digunakan tanpa kontrol atau tanpa bimbingan profesional. Penggunaan sistem klasifikasi ini harus disertai dengan kebijakan privasi yang ketat serta mekanisme pengamanan data untuk mencegah penyalahgunaan informasi yang dapat berdampak negatif bagi individu. Oleh karena itu, pengembangan teknologi ini harus selalu dibarengi dengan pertimbangan etis dan perlindungan hak privasi pengguna.

DAFTAR PUSTAKA

- [1] K. Yusrani, Ghefira, N. Aini, S. Maghfiroh, Aulia, and N. Istanti, Dwi, "Tinjauan Kebijakan Kesehatan Mental di Indonesia: Menuju Pencapaian Sustainable Development Goals dan Universal Health Coverage," *J. Med. Nusant.*, vol. 1, no. 2, pp. 89–107, 2023, doi: 10.59680/medika.v1i2.281.
- [2] C. Kokoh, H. Addarian, L. Anastasia, and D. Cahyadi, "Laporan Kasus: Gangguan Depresi Mayor Pada Mahasiswa Fakultas Kedokteran," *J. Pranata Biomedika*, vol. 3, no. 1, pp. 1–23, 2024.
- [3] H. Nathasya, P. Nuraini, S. Z. A. Thohiroh, T. Salma, and R. Fadhlina, "Analisis Tingkat Dan Faktor Penyebab Depresi Se Asia Tenggara," *J. Edu Res. Indones. Inst. Corp. Learn. Stud. Page*, vol. 15, no. 1, pp. 37–48, 2024.
- [4] A. Harahap, Juita, Sharmila, and Y. Mariska, "Pentingnya Menjaga Kesehatan Mental dalam Perspektif Agama Islam," *Innov. J. Soc. Sci. Res.*, vol. 4, no. 4, pp. 7836–7848, 2024.
- [5] E. J. Izati, R. Hairisyah, and E. Nurita, "Tantangan dan Solusi dalam Penanganan Kesehatan Mental di Kalangan Mahasiswa," *Pros. Semin. Nas. Manaj.*, vol. 4, no. 1, pp. 463–467, 2025.
- [6] A. Wijoyo, A. Y. Saputra, S. Ristanti, S. R. Sya'ban, M. Amalia, and R. Febriansyah, "Pembelajaran Machine Learning," *OKTAL (Jurnal Ilmu Komput. dan Sci.)*, vol. 3, no. 2, pp. 375–380, 2024, [Online]. Available: <https://journal.mediapublikasi.id/index.php/oktal/article/view/2305>
- [7] S. P. Utami, "Klasifikasi Kesehatan Mental Usia Remaja Menggunakan Algoritma Decision Tree Dan Naïve Bayes," Universitas Islam Negeri Syarif Hidayatullah, Jakarta, 2024.
- [8] A. P. Wijaya, "Perbandingan Algoritma Klasifikasi Random Forest dengan Naïve Bayes Classifier pada Studi Penyakit Berdasarkan Pola Nutrisi," *Remik Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 9, no. 1, pp. 429–438, 2025.
- [9] A. R. Dani and I. Handayani, "Klasifikasi Motif Batik Yogyakarta Menggunakan Metode GLCM Dan CNN," *J. Teknol. Terpadu*, vol. 10, no. 2, pp. 142–156, 2024.
- [10] Suci Amaliah, M. Nusrang, and A. Aswi, "Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 3, pp. 121–127, 2022, doi: 10.35580/variansiumn31.
- [11] R. Alfarez, E. Ermaitita, and R. M. B. Wadu, "Implementasi Algoritma Naïve Bayes Untuk Analisis Klasifikasi Survei Kesehatan Mental (Studi Kasus: Open Sourcing Mental Illness)," *Inform. J. Ilmu Komput.*, vol. 19, no. 1, pp. 1–10, 2023, doi: 10.52958/iftk.v19i1.4696.
- [12] A. Priyono, M. Shodiq, D. P. Alvinsyah, and S. A. Hidayah, "Metode Random Forest Untuk Memudahkan Klasifikasi Diagnosis Penyakit Mental," *J. Inform. Medis*, vol. 2, no. 1, pp. 1–4, 2024, doi: 10.52060/im.v2i1.2119.
- [13] A. A. Syam, G. H. M. A. Salim, D. F. Surianto, and M. F. B, "Analisis teknik preprocessing pada sentimen masyarakat terkait

- [14] Konflik israel-palestina menggunakan support vector machine,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 3, pp. 1464–1472, 2024.
- [15] D. Rifaldi, A. Fadlil, and Herman, “Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet ‘Mental Health.’” *Decod. J. Pendidik. Teknol. Inf.*, vol. 3, no. 2, pp. 161–171, 2023, doi: 10.51454/decode.v3i2.131.
- [16] D. Septiani and I. Isabela, “Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks,” *SINTEZIA J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 1, pp. 81–88, 2022.
- [17] R. Kristianto Hondro, “Jurnal Pendidikan Teknologi Informasi Dan Komputer Analisis Penerapan Text Mining dan TF-IDF dalam Mengetahui Sentimen Masyarakat Terhadap Kinerja POLRI,” *J. Pendidik. Teknol. Inf. Dan Komput.*, vol. 2, no. 1, pp. 44–49, 2023, [Online]. Available: <https://journal.grahamitra.id/index.php/petik>
- [18] R. Al Rasyid and D. H. U. Ningsih, “Penerapan Algoritma TF-IDF dan Cosine Similarity untuk Query Pencarian Pada Dataset Destinasi Wisata,” *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 8, no. 1, pp. 170–178, 2024, doi: 10.35870/jtik.v8i1.1416.
- [19] R. G. Wardhana, G. Wang, and F. Sibuea, “Penerapan Machine Learning Dalam Prediksi Tingkat Kasus Penyakit Di Indonesia,” *J. Inf. Syst. Manag.*, vol. 5, no. 1, pp. 40–45, 2023, doi: 10.24076/joism.2023v5i1.1136.
- [20] N. H. Setyawan and N. Wakhidah, “Analisis perbandingan metode logistic regression, random forest, gradient boosting untuk prediksi diabetes,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 1, pp. 150–162, 2025.
- [21] R. H. Kusumodestoni, M. Aan Presetyo, and A. Khanif Zyen, “Optimasi Algoritma Naïve Bayes Berbasis Kernel Untuk Klasifikasi Penyakit Hati,” *J. Inform. Teknol. dan Sains*, vol. 6, no. 3, pp. 748–756, 2024.
- [22] G. W. Nyipto Wibowo, S. Widiaastuti, M. Muratno, E. Lolang, and S. Soraya, “Penerapan Metode Teorema Bayes Dalam Mendiagnose Penyakit Tuberculosis,” *Build. Informatics, Technol. Sci.*, vol. 4, no. 4, pp. 1782–1788, 2023, doi: 10.47065/bits.v4i4.3035.
- [23] H. Hartono, A. Hajjah, and Y. N. Marlrim, “Penerapan Metode Naïve Bayes Classifier Untuk Klasifikasi Judul Berita,” *J. SimanteC*, vol. 12, no. 1, pp. 37–46, 2023.
- [24] S. Ary Prandika, D. P. Purba, P. Jojor Putri, and K. R. Bakara, “Implementasi Algoritma Random Forest Dalam Klasifikasi Diagnosis Penyakit Stroke,” *J. Penelit. Rumpun Ilmu Tek.*, vol. 2, no. 4, pp. 155–164, 2023, doi: 10.55606/juprit.v2i4.3039.
- [25] S. Mahmuda, “Implementasi Metode Random Forest pada Kategori Konten Kanal YouTube,” *J. Jendela Mat.*, vol. 2, no. 01, pp. 1–10, 2024, doi: 10.57008/jjm.v2i01.633.
- [26] R. Nurhidayat and K. E. Dewi, “Penerapan Algoritma K-Nearest Neighbor Dan Fitur Ekstraksi N-Gram Dalam Analisis Sentimen Berbasis Aspek,” *Komputa J. Ilm. Komput. dan Inform.*, vol. 12, no. 1, pp. 91–100, 2023, doi: 10.34010/komputa.v12i1.9458.
- [27] P. Romadloni, B. Adhi Kusuma, and W. Maulana Baihaqi, “Komparasi Metode Pembelajaran Mesin Untuk Implementasi Pengambilan Keputusan Dalam Menentukan Promosi Jabatan Karyawan,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6, no. 2, pp. 622–628, 2022, doi: 10.36040/jati.v6i2.5238.