Klasifikasi Gangguan Kecemasan Pengguna Twitter Menggunakan *Support Vector Machine*

1st Leinia Suryadi
Program Studi Teknologi Informasi
Universitas Telkom, Kampus Surabaya
Surabaya, Indonesia
leiniasuryadi@student.telkomuniversity.ac.id

2nd Bernadus Anggo Seno Aji
Program Studi Teknologi Informasi
Universitas Telkom, Kampus Surabaya
Surabaya, Indonesia
bernadusanggosenoaji@telkomuniversity.ac.id

3rd Mustafa Kamal
Program Studi Teknologi Informasi
Universitas Telkom, Kampus Surabaya
Surabaya, Indonesia
mustafakamal@telkomuniversity.ac.id

Abstrak — Gangguan mental yang umum adalah kecemasan, yang seringkali sulit terdeteksi karena tidak menunjukkan gejala fisik secara langsung serta dipengaruhi oleh rendahnya kesadaran masyarakat dan stigma negatif terhadap kesehatan jiwa. Akibatnya, banyak individu lebih memilih mengekspresikan perasaannya melalui media sosial seperti Twitter daripada mencari bantuan profesional. Namun, mendeteksi potensi gejala kecemasan melalui data teks bukanlah hal yang mudah karena pengguna jarang menyebutkan kondisi mentalnya secara eksplisit. Penelitian ini bertujuan merancang model klasifikasi gejala kecemasan pada pengguna Twitter menggunakan algoritma *Support Vector Machine* (SVM) dengan pendekatan *paraphrasing* berbasis IndoT5. Proses penelitian mencakup praproses teks dan pelatihan model SVM menggunakan kernel RBF dengan parameter optimal *C*=10 dan *gamma*=0,1. Hasil evaluasi menunjukkan bahwa penggunaan IndoT5 mampu meningkatkan performa model dengan capaian akurasi 97,52%, *precision* 97,57%, *recall* 97,50%, dan *F1-score* 97,52%. Dibandingkan algoritma *Multilayer Perceptron* (MLP) dan *Decision Tree*, SVM menunjukkan akurasi paling unggul. Model ini kemudian diimplementasikan ke sistem *web* berbasis Streamlit untuk mengklasifikasikan teks menjadi "Normal" atau "Kecemasan" sebagai alat bantu deteksi awal, bukan pengganti profesional.

Kata kunci — **kecemasan, twitter, klasifikasi teks, SVM, IndoT5.**

I. PENDAHULUAN

Gangguan mental merupakan suatu sindrom yang secara klinis ditandai oleh gangguan dalam pengendalian emosi atau perilaku, yang mencerminkan adanya disfungsi pada proses psikologis, biologis, atau perkembangan yang mendasari fungsi mental seseorang [1]. Gangguan mental yang paling umum adalah depresi dan kecemasan, diperkirakan mempengaruhi hampir 1 dari 10 orang (676 juta) [2]. Gangguan mental dapat menimbulkan penderitaan yang berdampak pada terganggunya kemampuan individu dalam menjalankan aktivitas sehari-hari [3].

Berbeda dengan gangguan fisik yang relatif mudah dikenali karena gejalanya tampak secara kasat mata, gangguan mental cenderung sulit terdeteksi karena tidak memperlihatkan tanda-tanda yang terlihat secara langsung. Kesulitan ini semakin diperparah oleh rendahnya tingkat kesadaran masyarakat akan pentingnya kesehatan jiwa, serta adanya stigma negatif terhadap individu yang mengalami gangguan mental. Akibatnya, banyak penderita merasa takut atau malu untuk mencari bantuan medis di fasilitas kesehatan [4]. Oleh karena itu, mereka beralih ke media sosial untuk mencari dukungan [5].

Twitter adalah salah satu saluran di mana individu dengan gangguan mental terhubung dengan jaringan yang lebih luas untuk mengungkapkan perasaan mereka, menjadikannya ruang untuk ekspresi emosi dan pencarian dukungan [6]. Mendeteksi individu yang berpotensi mengalami gangguan kecemasan melalui data berbasis teks bukanlah hal yang mudah, sebab mereka jarang mengungkapkan kondisi psikologisnya secara langsung [7]. Diperlukan suatu model yang mampu mengenali potensi kecemasan melalui teks, agar intervensi dini dapat dilakukan. Salah satu pendekatan yang dapat digunakan adalah teknik klasifikasi teks. Oleh karena itu, penelitian ini merancang sistem klasifikasi gejala kecemasan pada pengguna Twitter untuk mengenali potensi gangguan mental.

Dalam studi yang dilakukan oleh [8], analisis sentimen terhadap tingkat kepuasan pengguna layanan seluler di Indonesia melalui platform Twitter dengan menggunakan metode SVM menunjukkan akurasi sebesar 79% saat menerapkan fitur berbasis leksikon, dan meningkat menjadi 84% tanpa penggunaan fitur tersebut. Sementara itu, studi oleh [9], menggunakan metode SVM untuk analisis sentimen terhadap maskapai penerbangan melalui Twitter, menghasilkan akurasi sebesar 84,37% dengan kernel RBF, parameter complexity sebesar 1, dan gamma 1. Pada penelitian lain oleh [10], SVM diterapkan untuk mengidentifikasi kemungkinan adanya depresi dan kecemasan melalui data dari Twitter. Hasil klasifikasi terbaik diperoleh dengan kernel RBF, nilai C = 10, dan gamma = 0,1, yang memberikan akurasi sebesar 82,5%. Berdasarkan penelitian-penelitian sebelumnya, penelitian ini memilih metode *Support Vector Machine* (SVM) sebagai algoritma utama klasifikasi gejala kecemasan karena telah terbukti memiliki kinerja yang baik dalam pengolahan teks. Untuk mengatasi keragaman gaya bahasa pada *tweet*, penelitian ini juga menambahkan tahap praproses berupa *paraphrasing* menggunakan model IndoT5. Kombinasi pendekatan ini

menghasilkan model dengan tingkat akurasi sebesar 97,52% dalam mengenali gejala kecemasan pada pengguna Twitter.

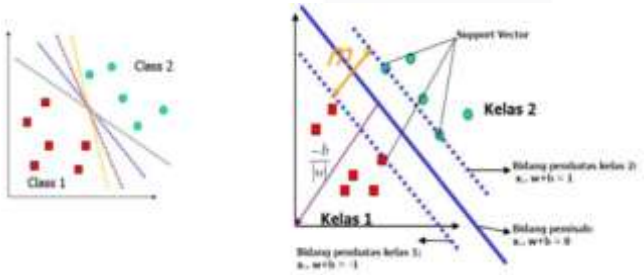
II. KAJIAN TEORI

A. Gangguan Kecemasan

Gangguan kecemasan *(anxiety disorders)* termasuk salah satu kategori gangguan mental yang dijelaskan dalam *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5). DSM-5 mendefinisikan gangguan kecemasan sebagai kondisi yang ditandai oleh rasa takut (*fear*) dan kekhawatiran (*anxiety*) yang berlebihan dan menetap, disertai perubahan perilaku terkait, sehingga mengganggu fungsi sehari-hari seseorang. Istilah "*fear*" merujuk pada reaksi emosional terhadap ancaman yang nyata, sedangkan "*anxiety*" menggambarkan kekhawatiran atas kemungkinan ancaman di masa depan. Perbedaan utama keduanya terletak pada pemicu dan ketahanannya: ketakutan bersifat lebih segera dan spesifik, sedangkan kecemasan lebih bertahan lama dan kadang tanpa pemicu yang jelas [11].

B. *Support Vector Machine* (SVM)

Support Vector Machine (SVM) merupakan salah satu algoritma dalam pembelajaran terawasi (*supervised learning*) yang digunakan untuk tugas klasifikasi. Prinsip kerja SVM adalah mencari *hyperplane* atau garis batas yang paling optimal dalam memisahkan dua kategori atau kelas data [12]. Gambar 1 memperlihatkan ilustrasi proses kerja metode ini.



GAMBAR 1
(HYPERPLANE SUPPORT VECTOR MACHINE (SVM) [13])

Untuk menentukan hyperplane yang paling optimal dalam memisahkan dua kelas data, dilakukan penghitungan margin dan pencarian titik maksimum. Persamaan (1) digunakan untuk memperoleh hyperplane dalam algoritma SVM.

$$(w.x_i) + b = 0 \quad (1)$$

Keterangan:
w: vektor bobot yang menentukan orientasi *hyperplane*,
x: vektor input (data) yang sedang diproses,
b: bias yang menentukan pergeseran *hyperplane* dari titik asal.

Di dalam data *xi*, yang termasuk kelas -1 dapat diperoleh melalui persamaan (2).

$$(w.x_i + b) \leq 1, y_i = -1 \quad (2)$$

Sedangkan data xi yang termasuk pada kelas +1 dapat dirumuskan seperti pada persamaan (3).

$$(w.x_i + b) \geq 1, y_i = 1 \quad (3)$$

Keterangan:
w: vektor bobot yang menentukan orientasi *hyperplane*
x: vektor input (data) yang sedang diproses
b: bias yang menentukan pergeseran *hyperplane* dari titik asal.

C. *K-Fold Cross Validation*

K-Fold Cross-Validation adalah teknik evaluasi yang membagi data menjadi *K* subset (*folds*), di mana proses pelatihan dan pengujian diulang sebanyak *K* kali agar setiap subset digunakan sebagai data uji satu kali. Metode ini bertujuan memberikan penilaian yang lebih akurat terhadap performa model SVM, mengurangi risiko *overfitting*, serta membantu memilih parameter terbaik untuk meningkatkan generalisasi model [14].

D. *Confusion Matrix*

Confusion Matrix alat yang digunakan untuk mengukur kinerja suatu model klasifikasi. Dalam *confusion matrix,* hasil prediksi dari suatu model dievaluasi dengan membandingkannya terhadap label kelas sebenarnya [15]. Data ini selanjutnya menjadi dasar dalam menghitung metrik evaluasi seperti akurasi, *presisi, recall,* dan *F1-score.* Contoh pengukuran dengan *confusion matrix* disajikan pada Tabel I.

TABEL 1
(CONFUSION MATRIX)

Data Aktual	**Data Prediksi**		
	TRUE	FALSE	TOTAL
TRUE	TP	FN	P
FALSE	FP	TN	N
TOTAL	P'	N'	P+N

Keterangan:
TP (True Positive) : Data positif yang terklasifikasi secara benar,
TN (True Negative) : Data negatif yang terklasifikasi secara benar,
FP (False Positive) : Data negatif yang terklasifikasi menjadi positif,
FN (False Negative) : Data positif yang terklasifikasi menjadi negatif.

Dari Tabel I tersebut, dapat dirumuskan persamaan yang digunakan untuk menghitung nilai akurasi, *precision, recall,* dan *F1-score* dari model klasifikasi.
Akurasi menunjukkan rasio antara jumlah prediksi yang tepat dengan jumlah keseluruhan data, dan dihitung menggunakan rumus pada persamaan (4).

$$Accuracy = \frac{TP + TN}{P + N} \quad (4)$$

Precision mengacu pada proporsi data positif yang diklasifikasikan dengan benar terhadap seluruh data yang diprediksi sebagai positif, sesuai dengan rumus (5).

$$Precision = \frac{TP}{P'} \quad (5)$$

Recall adalah proporsi data berlabel positif yang berhasil diklasifikasikan dengan benar dibandingkan dengan total data positif aktual, sebagaimana dirumuskan dalam persamaan (6).

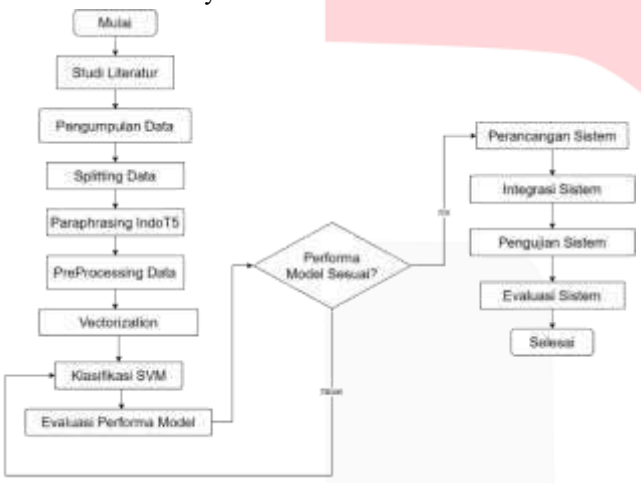
$$Recall = \frac{TP}{P} \quad (6)$$

F1 score merupakan rata-rata harmonis dari precision dan recall, dan dihitung menggunakan persamaan yang tercantum pada rumus (7).

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (7)$$

III. METODE

A. Sistematika Penyelesaian Masalah



GAMBAR 2
(ALUR PROSEDUR PENELITIAN)

B. Studi Literatur

Studi literatur dilakukan sebagai langkah awal untuk mengidentifikasi permasalahan penelitian serta meninjau pendekatan dan solusi yang telah ditawarkan dalam penelitian sebelumnya guna menentukan metode yang paling sesuai.

C. Pengumpulan Data

Penelitian ini menggunakan data sekunder yang diperoleh dari platform Kaggle, yakni dataset berjudul *Depression and Anxiety in Twitter (ID)* yang disusun oleh Steven Hans. Dataset ini terdiri dari dua komponen, yaitu teks dan label. Pada label terdapat dua, yaitu 0 untuk mendefinisikan Normal dan 1 untuk mendefinisikan Gejala Kecemasan. Total *tweet* pada dataset ini berjumlah 6980.

D. *Splitting Data*

Dataset yang telah dikumpulkan dibagi menjadi dua subset: data latih (*training data*) dan data uji (*testing data*). Dilakukan percobaan dengan berbagai rasio pembagian data, yaitu 80:20, 70:30, dan 60:40. Berikut adalah hasil akurasi untuk masing-masing pembagian data menggunakan kernel *linear* tanpa parameter:

- Rasio 80:20 menghasilkan akurasi sebesar **99,85%**
- Rasio 70:30 menghasilkan akurasi sebesar **99,90%**
- Rasio 60:40 menghasilkan akurasi sebesar **99,78%**

Dari hasil tersebut, rasio 70:30 dipilih sebagai konfigurasi terbaik karena memberikan akurasi tertinggi sekaligus menjaga keseimbangan antara data latih (sebanyak 4.886 data) dan data uji (2.094 data).

E. *Text-to-Text Transfer Transformer* (IndoT5)

IndoT5 merupakan adaptasi dari model *Text-to-Text Transfer Transformer* (T5) yang dioptimalkan untuk tugas pemrosesan bahasa alami (NLP) dalam konteks bahasa Indonesia. Dengan pendekatan *sequence-to-sequence,* model ini terbukti efektif untuk menghasilkan *parafrase*, ringkasan, dan frasa kunci, yang menjaga kesetaraan semantik namun dengan struktur berbeda suatu karakteristik penting dalam *augmentasi* data untuk klasifikasi teks [16].

Kelebihan IndoT5 terletak pada kemampuannya menghasilkan *parafrase* yang koheren dan sesuai konteks, terutama setelah melalui proses *fine-tuning* pada korpus bahasa Indonesia. Model ini juga mendukung generasi multi-variasi dari satu input, memberikan kontribusi signifikan terhadap peningkatan representasi data minoritas. Evaluasi model menunjukkan performa tinggi berdasarkan skor *ROUGE* dan *BLEU*, menandakan kualitas semantik dan sintaktik parafrase yang dihasilkan tetap terjaga [17].

Dalam penelitian ini, kerangka CRISP-DM *(Cross-Industry Standard Process for Data Mining)* digunakan untuk membimbing proses analisis data secara sistematis. Kerangka ini dipilih karena bersifat netral terhadap industri serta terbukti efektif dan fleksibel dalam berbagai penerapan *data mining* [18].

GAMBAR 3
(ALUR CRISP-DM INDOT5)

a) *Data Preparation*

Pada tahap ini dilakukan dengan menyeleksi data berdasarkan kelengkapan dan validitasnya, serta memastikan distribusi label yang seimbang untuk mendukung proses pelatihan model. Proses ini mencakup penghapusan data kosong, pengelompokan data berdasarkan label klasifikasi, dan penambahan variasi teks menggunakan teknik augmentasi berbasis *paraphrasing* agar data minoritas dapat ditingkatkan jumlahnya secara signifikan. Hasil akhir dari tahap ini adalah kumpulan data bersih dan seimbang yang siap digunakan dalam pemodelan klasifikasi.

b) *Modelling*

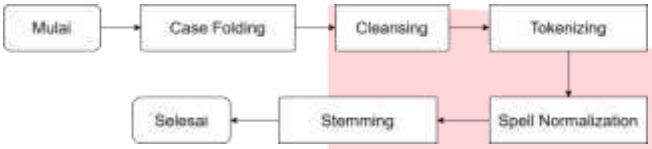
Tahapan ini menggunakan model IndoT5 *(Text-to-Text Transfer Transformer)* untuk membangun sistem parafrase otomatis berbasis data yang telah diproses. Model ini dilatih untuk memahami input teks dalam bahasa Indonesia dan menghasilkan keluaran berupa

parafrase, dengan tetap mempertahankan makna yang sama. IndoT5 bekerja dengan pendekatan *sequence-to-sequence* untuk menyusun kembali kalimat dalam bentuk yang berbeda namun tetap konsisten secara semantik.

c) Evaluasi Model

Melakukan evaluasi model menggunakan metrik *BLEU, ROUGE,* dan *METEOR* untuk mengukur kinerja algoritma IndoT5 *(Text-to-Text Transfer Transformer)*, untuk menilai sejauh mana keluaran model dapat mempertahankan makna dan kualitas teks dibandingkan dengan referensi yang tersedia.

F. *Pre-Processing*



GAMBAR 4
(ALUR PRE-PROCESSING)

Proses awal dalam pengolahan data adalah tahap *preprocessing*, yakni serangkaian langkah untuk membersihkan dan menstandarkan data teks sebelum dianalisis lebih lanjut. Rincian metode yang digunakan dalam tahap ini sebagaimana dijelaskan dalam [9], [10] meliputi:

a) Case Folding

Merupakan proses normalisasi teks dengan mengubah seluruh huruf menjadi huruf kecil (*lowercase*) dan menghilangkan karakter-karakter yang tidak diperlukan seperti angka maupun tanda baca.

b) Cleansing

Tahapan ini mencakup pembersihan teks dari elemen-elemen yang tidak relevan atau mengganggu (*noise*), seperti tag HTML, emotikon, tanda pagar (#), nama pengguna (*@username*), *retweet* (RT), serta tautan URL atau alamat situs.

c) Tokenizing

Tokenisasi adalah proses memecah teks menjadi unit-unit kata (token) yang dapat dianalisis secara individual.

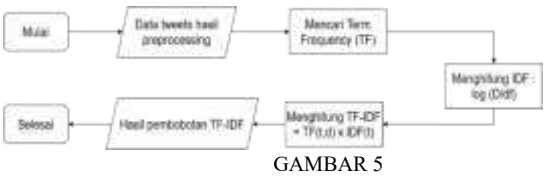
d) Spell Normalization

Bertujuan mengonversi kata tidak baku atau kata slang ke dalam bentuk baku untuk mengurangi variasi kata yang dapat mengganggu akurasi model.

e) Stemming

Stemming dilakukan dengan mengubah kata berimbuhan menjadi bentuk dasarnya sesuai dengan kaidah morfologi. Penelitian ini memanfaatkan *library* Sastrawi yang dikenal memiliki akurasi tinggi dalam pemrosesan morfologis Bahasa Indonesia [19].

G. Vectorization



GAMBAR 5
(ALUR VECTORIZATION)

Pada tahap ini, kumpulan kata atau istilah dalam dokumen akan diubah menjadi representasi numerik agar dapat diolah dalam proses klasifikasi. Pembobotan yang digunakan dalam penelitian ini adalah metode *Term Frequency-Inverse Document Frequency* (TF-IDF).

H. Klasifikasi *Support Vector Machine* (SVM)

Model SVM memanfaatkan beberapa parameter utama yang disesuaikan untuk memaksimalkan performa klasifikasi gejala kecemasan. Parameter tersebut dapat dilihat pada Tabel 2.

TABEL 2
(KERNEL DAN PARAMETER PILIHAN)

No.	Parameter	Keterangan
1.	Kernel	Jenis kernel yang digunakan dalam SVM. Kernel yang dipertimbangkan adalah *linear, polynomial, sigmoid*, dan RBF.
2.	*C*	Parameter regulasi yang menentukan keseimbangan antara margin maksimal dan kesalahan klasifikasi.
3.	*Gamma*	Parameter yang mengatur pengaruh satu data latih pada keseluruhan model.
4.	*Degree*	parameter yang mengontrol derajat polinomial yang digunakan dalam kernel.

Proses *hyperparameter tuning* dilakukan menggunakan *GridSearchCV* untuk menentukan kombinasi parameter terbaik yang dapat meningkatkan kinerja klasifikasi SVM.

TABEL 3
(KOMBINASI NILAI PARAMETER SVM)

Kernel	Parameter	Nilai
Linear	*Complexity (C)*	[0.001, 0.01, 0.1, 1, 5, 10, 50, 100]
RBF	*Complexity (C)*	[0.001, 0.01, 0.1, 1, 10, 100, 500]
	Gamma (γ)	[0.0001, 0.001, 0.01, 0.1, 1, 10]
Polynomial	*Degree (d)*	[1, 2, 3, 4]
	Complexity (C)	[0.1, 1, 10, 50, 100]
Sigmoid	*Gamma (γ)*	[0.0001, 0.001, 0.01, 0.1, 1]
	Complexity (C)	[0.1, 1, 10, 100, 500]

Proses *hyperparameter tuning* dilakukan bersamaan dengan *K-Fold Cross-Validation*, di mana dataset dibagi menjadi K lipatan *(fold)*. Dalam penelitian ini, digunakan *K=5* sebagai jumlah *iterasi* untuk *cross-validation*. Teknik ini bertujuan untuk mengevaluasi performa model secara menyeluruh, meminimalkan bias, dan mencegah *overfitting*.

I. Evaluasi Model

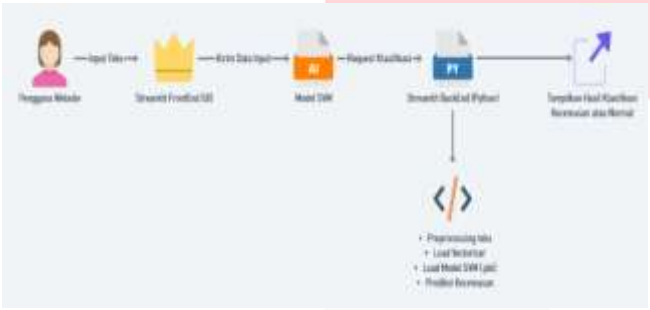
Evaluasi model dilakukan menggunakan *Confusion Matrix* untuk menghitung metrik *accuracy*, *precision*, *recall*, dan *f1-score*. *Recall* menjadi prioritas dalam mendeteksi gejala untuk meminimalkan *false negative*, sedangkan *precision* penting untuk menghindari *false positive*. *F1-score* memberikan keseimbangan antara keduanya, khususnya pada data tidak seimbang. Selain SVM, penelitian ini juga membandingkan performa dengan algoritma *Multilayer Perceptron* (MLP) dan

Decision Tree menggunakan dataset dan parameter yang sama [20].

TABEL 4
(ALGORITMA PEMBANDING SVM)

Kernel	Parameter dan Nilai
MLP *(Multilayer Perceptron)*	hidden_layer_sizes: (25,)
	solver: 'adam'
	learning_rate_init: 1e-3
	max_iter: 100
	random_state: 42
Decision Tree	criterion: 'gini'
	min_samples_split: 2
	min_samples_leaf: 1
	class_weight: 'balanced'
	random_state: 42

J. Arsitektur Sistem



GAMBAR 6
(ARSITEKTUR SISTEM)

Penelitian ini merancang sistem klasifikasi gejala kecemasan berbasis model SVM yang diintegrasikan ke dalam aplikasi *web* menggunakan Streamlit. Pengguna memasukkan teks melalui antarmuka aplikasi, kemudian sistem *backend* memproses teks tersebut melalui tahapan *preprocessing*, ekstraksi fitur dengan TF-IDF, dan klasifikasi menggunakan model SVM terlatih. Hasil prediksi ditampilkan kembali ke pengguna melalui antarmuka Streamlit. Seluruh proses berlangsung secara otomatis dan terintegrasi dalam satu sistem berbasis Python.

IV. HASIL DAN PEMBAHASAN

A. Evaluasi *Support Vector Machine* (SVM)

Model *Support Vector Machine* (SVM) dipilih sebagai algoritma klasifikasi utama dalam penelitian ini. Pengembangan model dilakukan melalui proses penyetelan *hyperparameter* menggunakan *GridSearchCV* dengan skema validasi silang 5-*fold*. Konfigurasi terbaik diperoleh pada *iterasi* ke-39 dengan parameter kernel RBF**,** *C*=10, dan *gamma*=0.1**,** yang menghasilkan rata-rata akurasi validasi tertinggi sebesar 0,9659 (96,59%)**.** Skor akurasi pada masing-masing *fold* berkisar antara 0,9599 hingga 0,9742.

Evaluasi model SVM dengan kernel RBF dan *paraphrasing* IndoT5 pada 2094 data uji menunjukkan performa klasifikasi yang tinggi. Sebanyak 1842 data Normal dan 200 data Kecemasan berhasil diklasifikasikan dengan benar. Kesalahan prediksi tercatat 32 kasus *false positive* dan 20 kasus *false negative*. Hasil ini menunjukkan bahwa model mampu membedakan teks normal dan teks dengan indikasi kecemasan secara akurat, dengan tingkat kesalahan yang rendah.

TABEL 5
(CONFUSION MATRIX SVM DENGAN INDOT5)

		Actual Values	
		Normal	Kecemasan
Predicted Values	Normal	1842	32
	Kecemasan	20	200

Evaluasi performa model SVM dilakukan berdasarkan metrik standar yang mencakup *precision, recall,* dan *F1-score* untuk masing-masing kelas pada data uji sebanyak 2094 sampel. Model dikembangkan dengan pendekatan *augmentasi* data menggunakan IndoT5 guna meningkatkan variasi dan keseimbangan antar kelas.

TABEL 6
(METRIK EVALUASI PER KELAS MODEL SVM DENGAN INDOT5)

Kelas	Precision	Recall	F1-score
Normal	0.8621	0.9091	0.8850
Kecemasan	0.9893	0.9829	0.9861

Untuk kelas Normal, model memperoleh *precision* sebesar 0,8621, *recall* sebesar 0,9091, dan *F1-score* sebesar 0,8850, mencerminkan ketepatan dan cakupan klasifikasi yang cukup baik. Sementara itu, pada kelas Kecemasan, model menunjukkan performa yang lebih tinggi dengan *precision* sebesar 0,9893, *recall* sebesar 0,9829, dan *F1-score* sebesar 0,9861.

Hasil ini mengindikasikan bahwa model sangat andal dalam mengenali gejala kecemasan dengan kesalahan prediksi yang sangat rendah. Kinerja keseluruhan model juga kuat, dengan total prediksi benar sebesar 2042 dari 2094 data uji, menghasilkan akurasi sebesar 97,5%. Nilai rata-rata tertimbang *(weighted average)* untuk *precision*, *recall*, dan *F1-score* di kisaran 97%, menunjukkan performa klasifikasi yang konsisten dan seimbang di seluruh kelas setelah penerapan teknik *augmentasi* IndoT5.

Selanjutnya, Pengembangan model tanpa *paraphrasing* dilakukan dengan penyetelan *hyperparameter* menggunakan *GridSearchCV* dan skema validasi silang 5-*fold cross-validation.* Konfigurasi terbaik diperoleh pada *iterasi* ke-88, dengan parameter kernel *sigmoid*, *C* = 1, dan *gamma* = 1. Konfigurasi ini menghasilkan rata-rata akurasi validasi tertinggi sebesar 0,9928 (99,28%), Skor akurasi pada masing-masing *fold* berkisar antara 0,9857 hingga 0,9969.

Evaluasi lanjutan dilakukan terhadap model SVM menggunakan dataset asli tanpa *augmentasi paraphrasing* IndoT5. Berdasarkan hasil *confusion matrix* terhadap 2094 data uji, model mampu mengklasifikasikan 1874 data Normal dan 205 data Kecemasan secara benar. Namun, terdapat 15 data Kecemasan yang salah diklasifikasikan sebagai Normal, menunjukkan adanya kecenderungan bias terhadap kelas mayoritas (Normal). Tidak ditemukannya kesalahan klasifikasi untuk kelas Normal (*false positive* = 0) mengindikasikan sensitivitas model terhadap distribusi data yang tidak seimbang.

TABEL 7
(CONFUSION MATRIX SVM TANPA INDOT5)

		Actual Values	
		Normal	Kecemasan
Predicted Values	Normal	1874	0
	Kecemasan	15	205

Evaluasi performa model SVM dilakukan berdasarkan metrik standar yang mencakup *precision, recall,* dan *F1-score* untuk masing-masing kelas pada data uji sebanyak 2094 sampel. Berbeda dengan pendekatan sebelumnya, model ini dibangun tanpa menggunakan teknik *augmentasi* data IndoT5, sehingga hanya mengandalkan variasi alami dari dataset asli.

TABEL 8
(METRIK EVALUASI PER KELAS MODEL SVM TANPA INDOT5)

Kelas	Precision	Recall	F1-score
Normal	1.0000	0.9318	0.9647
Kecemasan	0.9921	1.0000	0.9960

Untuk kelas Normal, model memperoleh *precision* sempurna sebesar 1,0000, namun dengan *recall* sebesar 0,9318, yang berarti masih terdapat sejumlah data Normal yang gagal dikenali secara tepat. Nilai *F1-score* untuk kelas ini tercatat sebesar 0,9647, mencerminkan ketidakseimbangan antara ketepatan dan cakupan klasifikasi. Sementara itu, pada kelas Kecemasan, model menunjukkan performa yang sangat tinggi dengan *precision* sebesar 0,9921 dan *recall* sempurna sebesar 1,0000, menghasilkan F1-score sebesar 0,9960.

Hasil ini mengindikasikan bahwa model memiliki kemampuan klasifikasi yang kuat secara umum, namun performanya cenderung bias terhadap kelas mayoritas, yaitu kelas Normal. Akurasi tinggi yang dicapai, yaitu 99,28%, tidak sepenuhnya mencerminkan kemampuan generalisasi model yang merata. Nilai rata-rata tertimbang *(weighted average)* untuk *precision, recall,* dan *F1-score* berada di kisaran 99,2%, namun perbedaan *recall* antara kedua kelas menunjukkan bahwa model kurang optimal dalam mengenali variasi data pada kelas Kecemasan. Hal ini memperkuat pentingnya teknik *augmentasi* seperti IndoT5 untuk meningkatkan keseimbangan model dalam mendeteksi kelas minoritas secara lebih adil dan akurat.

TABEL 9
(TABEL PERBANDINGAN HASIL EVALUASI MODEL SVM)

Model	Accuracy	Precision	Recall	F1-Score
SVM dengan IndoT5	**0.9751**	**0.9757**	**0.9950**	**0.9752**
SVM tanpa IndoT5	0.9928	0.9928	0.9928	0.9927

Meskipun model tanpa IndoT5 menunjukkan nilai metrik lebih tinggi, hasilnya cenderung bias terhadap kelas mayoritas. Penggunaan IndoT5 dalam penelitian ini terbukti meningkatkan keseimbangan data, memperbaiki generalisasi model, dan memungkinkan klasifikasi yang lebih adil terhadap kedua kelas.

B. Evaluasi Model Pembanding

Untuk menguji keandalan model SVM dan menilai efektivitas pendekatan yang digunakan, dilakukan evaluasi terhadap dua algoritma pembanding, yaitu *Multilayer Perceptron* (MLP) dan *Decision Tree*. Kedua model pembanding ini diuji menggunakan data uji yang sama agar hasil perbandingan bersifat adil dan konsisten.

a) Evaluasi Model *Multilayer Perceptron* (MLP)

TABEL 10
(CONFUSION MATRIX MULTILAYER PERCEPTRON (MLP))

		Actual Values	
		Normal	Kecemasan
Predicted Values	Normal	1789	85
	Kecemasan	53	167

Berdasarkan *confusion matrix* pada model *Multilayer Perceptron* (MLP), dari total 1.874 data uji untuk kelas Normal, model berhasil memprediksi dengan benar sebanyak 1.789 data dan salah memprediksi sebanyak 85 data. Untuk kelas Kecemasan yang berjumlah 220 data, model berhasil mengklasifikasikan dengan benar sebanyak 167 data dan salah klasifikasi sebanyak 53 data. Hasil ini menunjukkan bahwa model MLP mampu mengenali pola data dengan cukup baik, tetapi masih terdapat kesalahan prediksi yang lebih tinggi pada kelas Kecemasan dibandingkan pada kelas Normal. Nilai ini menunjukkan bahwa model MLP lebih baik dalam mengenali kelas Kecemasan, tetapi performanya kurang optimal pada kelas Normal karena masih banyak terjadi kesalahan prediksi

TABEL 11
(METRIK EVALUASI PER KELAS MODEL MLP)

Kelas	Precision	Recall	F1-score
Normal	0.6627	0.7591	0.7076
Kecemasan	0.9712	0.9546	0.9629

Hasil evaluasi model *Multilayer Perceptron* (MLP) menunjukkan bahwa pada kelas Normal diperoleh *precision* sebesar 0,6627, *recall* sebesar 0,7591, dan *F1-score* sebesar 0,7076. Sementara itu, pada kelas Kecemasan, model menghasilkan *precision* sebesar 0,9712, *recall* sebesar 0,9546, dan *F1-score* sebesar 0,9629.

Berdasarkan hasil evaluasi model *Multilayer Perceptron* (MLP) terhadap data uji, diperoleh akurasi sebesar 93,42%. Pada kelas Normal, model memiliki *precision* sebesar 0,9712, *recall* sebesar 0,9546, dan *F1-score* sebesar 0,9629. Sementara itu, pada kelas Kecemasan, *precision* yang dihasilkan sebesar 0,6627, *recall* sebesar 0,7591, dan *F1-score* sebesar 0,7076. Nilai rata-rata tertimbang *(weighted average)* untuk *precision*, *recall,* dan *F1-score* masing-masing sebesar 0,9392; 0,9342; dan 0,9361. Hasil ini menunjukkan bahwa model MLP memiliki performa yang cukup baik, namun masih kurang seimbang karena kesalahan prediksi lebih banyak terjadi pada kelas minoritas (Kecemasan) dibandingkan kelas mayoritas (Normal).

b) Evaluasi Model Decision Tree

TABEL 12
(CONFUSION MATRIX DECISION TREE)

		Actual Values	
		Normal	Kecemasan
Predicted Values	Normal	1798	76
	Kecemasan	7	213

Berdasarkan *confusion matrix* pada model *Decision Tree,* dari total 1874 data uji untuk kelas Normal, model berhasil memprediksi dengan benar sebanyak 1798 data dan salah memprediksi sebanyak 76 data. Untuk kelas Kecemasan yang berjumlah 220 data, model berhasil diklasifikasikan dengan benar sebanyak 213 data dan salah diprediksi sebagai kelas Normal sebanyak 7 data. Hasil ini menunjukkan bahwa model *Decision Tree* mampu mengenali kedua kelas dengan cukup baik, walaupun masih terdapat kesalahan prediksi pada kelas Normal yang terdeteksi sebagai Kecemasan.

TABEL 13
(METRIK EVALUASI PER KELAS MODEL DECISION TREE)

Kelas	Precision	Recall	F1-score
Normal	0.7370	0.9682	0.8369
Kecemasan	0.9961	0.9594	0.9774

Hasil evaluasi model *Decision Tree* menunjukkan bahwa pada kelas Normal diperoleh *precision* sebesar 0,7370, *recall* sebesar 0,9682, dan *F1-score* sebesar 0,8369. Sementara itu, pada kelas Kecemasan, model menghasilkan *precision* sebesar 0,9961, *recall* sebesar 0,9594, dan *F1-score* sebesar 0,9774.

Berdasarkan hasil evaluasi, model *Decision Tree* mencapai tingkat akurasi sebesar 96,04%. Pada kelas Normal, *precision* yang diperoleh adalah 0,9961 dengan *recall* sebesar 0,9594 sehingga menghasilkan *F1-score* 0,9774. Sementara itu, pada kelas Kecemasan, model menghasilkan *precision* sebesar 0,7370, *recall* sebesar 0,9682, dan *F1-score* sebesar 0,8369. Hasil *weighted average* untuk *precision, recall,* dan *F1-score* masing-masing adalah 0,9688; 0,9600; dan 0,9630. Nilai ini menunjukkan bahwa model *Decision Tree* mampu mengenali data dengan cukup baik, terutama pada kelas Normal, namun performa untuk kelas Kecemasan masih terbatas karena *precision* yang lebih rendah.

C. Perbandingan Keseluruhan Model

TABEL 14
(PERBANDINGAN KESELURUHAN MODEL)

Model	Accuracy	Precision	Recall	F1-Score
Support Vector Machine (SVM) dengan IndoT5	**0.9751**	**0.9757**	**0.9750**	**0.9752**
Multilayer Perceptron (MLP) dengan IndoT5	0.9342	0.9392	0.9342	0.9361
Decision Tree dengan IndoT5	0.9604	0.9688	0.9600	0.9630
Support Vector Machine (SVM) tanpa IndoT5	0.9928	0.9928	0.9928	0.9927
Multilayer Perceptron (MLP) tanpa IndoT5	0.9679	0.9675	0.9661	0.9658
Decision Tree tanpa IndoT5	0.9938	0.9830	0.9625	0.9727

Meskipun SVM tanpa IndoT5 menghasilkan akurasi paling tinggi, model ini cenderung bias ke kelas mayoritas (Normal) dan berisiko *overfitting*. Ini terlihat dari kemampuannya yang sangat baik dalam mengenali data Normal, tapi masih gagal mendeteksi sebagian data Kecemasan. Dalam kasus deteksi gejala kecemasan, mengenali kelas minoritas justru lebih penting. Karena itu, SVM dengan IndoT5 bisa dianggap lebih seimbang. Meskipun akurasinya sedikit lebih rendah, model ini lebih baik dalam mengenali data Kecemasan. Proses *augmentasi* dengan IndoT5 membantu memperkaya variasi data, sehingga model bisa belajar dari pola yang lebih beragam. Ini membuat SVM dengan IndoT5 lebih cocok digunakan, meskipun tidak semua model mendapat peningkatan signifikan dari IndoT5.

D. Analisis Data Salah Prediksi

Hasil analisis kesalahan menunjukkan bahwa model SVM dengan kernel RBF memiliki jumlah kesalahan prediksi paling sedikit (52 kasus), sebagian besar disebabkan oleh teks ambigu yang sulit dibedakan konteks emosinya. *Decision Tree* mencatat 83 kesalahan, dengan kecenderungan salah mengklasifikasikan teks netral sebagai kecemasan akibat kelemahan dalam memahami konteks. Sementara itu, MLP menunjukkan performa terburuk dengan 138 kesalahan, yang tersebar merata pada kedua kelas dan mencerminkan kesulitan dalam menangani data kompleks dan tidak seimbang. Secara keseluruhan, analisis ini menegaskan bahwa SVM merupakan model paling andal dalam penelitian ini, baik dari sisi jumlah kesalahan maupun hasil evaluasi metrik klasifikasi.

TABEL 15
(CONTOH ANALISIS SALAH PREDIKSI MODEL SVM)

Index	Teks	Label	Prediksi
37	warning kapel dapat sebab kanker serang jantung impotensi dan ganggu hamil dan janin	Normal	Kecemasan
109	tidak nikmat banget dicuekin	Normal	Kecemasan
112	yaallah nilai aku udh keluar semua takut banget	Normal	Kecemasan
138	kenapa semua berahir seperti ini lagi	Normal	Kecemasan
242	se cemas itu aku nanya sama diri aku yang lain sampai kirim rezeki punya temen yang baik haru	Kecemasan	Normal

E. Hasil Pengujian Sistem

Pengujian sistem dilakukan melalui antarmuka aplikasi berbasis Streamlit dengan memberikan input teks secara langsung. Pada gambar pertama, sistem menerima teks yang berisi keluhan tentang rasa takut bertemu orang, gejala *nervous* berlebihan, dan keringat berlebih. Sistem berhasil mengklasifikasikan teks tersebut ke dalam kategori Kecemasan, sesuai dengan konteks isi teks.

GAMBAR 7
(PENGUJIAN SISTEM PADA HASIL PREDIKSI KECEMASAN)

Pada gambar kedua, input teks berupa ekspresi kebahagiaan karena mendengar berita tentang grup musik favorit. Sistem dengan tepat mengklasifikasikan teks tersebut ke dalam kategori Normal, karena tidak menunjukkan adanya indikasi kecemasan. Hasil pengujian ini menunjukkan bahwa sistem dapat memberikan prediksi yang sesuai dengan konteks teks masukan, baik untuk kategori Kecemasan maupun Normal.

GAMBAR 8
(PENGUJIAN SISTEM PADA HASIL PREDIKSI NORMAL)

V. KESIMPULAN

Penelitian ini berhasil merancang model klasifikasi gejala kecemasan pengguna Twitter menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel RBF, di mana proses optimasi *hyperparameter* menghasilkan konfigurasi terbaik pada C = 10 dan *gamma* = 0,1. Penggunaan teknik *paraphrasing* dengan IndoT5 terbukti meningkatkan keseimbangan klasifikasi, menghasilkan akurasi 97,51% dengan performa *precision, recall,* dan *F1-score* yang seimbang, serta mengurangi bias terhadap kelas minoritas. SVM dengan IndoT5 menunjukkan kinerja lebih baik dibandingkan *Multilayer Perceptron* dan *Decision Tree.* Model ini kemudian berhasil diimplementasikan dalam sistem berbasis *web* menggunakan Streamlit, yang mampu mengklasifikasikan teks secara akurat sebagai Normal atau Kecemasan sesuai dengan isi *tweet* pengguna.

REFERENSI

[1] V. del Barrio, "Diagnostic and Statistical Manual of Mental Disorders," in *Encyclopedia of Applied Psychology*, C. D. Spielberger, Ed. Elsevier, 2004, pp. 607–614.

[2] World Health Organization, "WORLD HEALTH STATISTICS - MONITORING HEALTH FOR THE SDGs," *World Heal. Organ.*, p. 1.121, 2016.

[3] D. Bolton, *What is Mental Disorder? An essay in philosophy, science, and values*. Oxford University Press, 2008.

[4] A. H. Orabi, P. Buddhitha, M. H. Orabi, and D. Inkpen, "Deep learning for depression detection of Twitter users," *Proc. 5th Work. Comput. Linguist. Clin. Psychol. From Keyboard to Clin. CLPsych 2018 2018 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol.*, pp. 88–97, 2018, doi: 10.18653/v1/w18-0609.

[5] A. Yates, A. Cohan, and N. Goharian, "Depression and self-harm risk assessment in online forums," *EMNLP 2017 - Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 2968–2978, 2017, doi: 10.18653/v1/d17-1322.

[6] K. Yan, D. Arisandi, and T. Tony, "Analisis Sentimen Komentar Netizen Twitter Terhadap Kesehatan Mental Masyarakat Indonesia," *J. Ilmu Komput. dan Sist. Inf.*, vol. 10, no. 1, 2022, doi: 10.24912/jiksi.v10i1.17865.

[7] D. Owen, J. C. Collados, and L. Espinosa-Anke, "Towards Preemptive Detection of Depression and Anxiety in Twitter," pp. 82–89, 2020, [Online]. Available: http://arxiv.org/abs/2011.05249.

[8] U. Rofiqoh, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1725–1732, 2017, [Online]. Available: http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628.

[9] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, 2021, doi: 10.34148/teknika.v10i1.311.

[10] F. Darmawan, M. Joe, Y. I. Kurniawan, and L. Afuan, "Analisis Sentimen Kemungkinan Depresi dan Kecemasan pada Twitter Menggunakan Support Vector Machine," *J. Eksplora Inform.*, vol. 13, no. 1, pp. 24–36, 2023, doi: 10.30864/eksplora.v13i1.854.

[11] V. del Barrio, *Diagnostic and Statistical Manual of Mental Disorders*, vol. 1. 2004.

[12] S. Lidya, S.K., Sitompul, O.S., & Efendi, "Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine (SVM) dan K-Nearest Neighbor (K-NN)," 2015. [Online]. Available: http://repositori.usu.ac.id/handle/123456789/42877.

[13] K. Sembiring, "Penerapan Teknik Support Vector Machine untuk Pendeteksian Intrusi pada Jaringan.," no. September, pp. 1–28, 2007.

[14] L. Qadrini *et al.*, "Ensemble Resampling Support Vector Machine, Multinomial Regression To Multiclass Imbalanced Data," *BAREKENG J. Ilmu Mat. Dan Terap.*, vol. 18, no. 1, pp. 0269–0280, 2024, doi: 10.30598/barekengvol18iss1pp0269-0280.

[15] C. I. Salekhah, "Implementasi Metode Multi Class Support Vector Machine Untuk Klasifikasi Emosi Pada Lirik Lagu Bahasa Indonesia," Universitas Komputer Indonesia (Unikom), 2016.

[16] U. Saokani, M. Irfan, D. S. Maylawati, R. J. Abidin, I. Taufik, and R. N. Hay's, "Comparison of the Fisher-Yates Shuffle and the Linear Congruent Algorithm for Randomizing Questions in Nahwu Learning Multimedia," *Khazanah J. Relig. Technol.*, vol. 1, no. 1, pp. 10–14, 2023, doi: 10.15575/kjrt.v1i1.159.

[17] W. Puspitasari, "IndoT5 (Text-to-Text Transfer Transformer) Algorithm for Paraphrasing Indonesian Language Islamic Sermon Manuscripts," vol. 2, no. 2, 2025.

[18] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.

[19] D. A. N. Arifin, S. Pada, D. Teks, B. Indonesia, J. Pardede, and D. Darmawan, "PERBANDINGAN ALGORITMA STEMMING PORTER , SASTRAWI , IDRIS , COMPARISON OF STEMMING ALGORITHMS PORTER , SASTRAWI , IDRIS , AND ARIFIN SETIONO ON INDONESIAN TEXT DOCUMENTS," vol. 12, no. 1, 2025, doi: 10.25126/jtiik.2025128860.

[20] K. S. Nugroho, I. Akbar, A. N. Suksmawati, and Istiadi, "Deteksi Depresi dan Kecemasan Pengguna Twitter Menggunakan Bidirectional LSTM," no. Ciastech, pp. 287–296, 2023, [Online]. Available: http://arxiv.org/abs/2301.04521.