

# Deteksi Stres Berbasis Teks pada Dreddit Menggunakan Fine Tuning DeBERTa-v3

Pramudia Ramadhan\*, De Rosal Ignatius Moses Setiadi

Fakultas Ilmu dan Komputer, Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

Email: <sup>1</sup>\*111202214274@mhs.dinus.ac.id, <sup>2</sup>moses@dsn.dinus.ac.id

Email Penulis Korespondensi: 111202214274@mhs.dinus.ac.id

Submitted 06-11-2025; Accepted 01-12-2025; Published 15-12-2025

## Abstrak

Kesehatan mental menjadi isu penting di era digital karena ekspresi psikologis individu banyak terekam melalui unggahan media sosial seperti Reddit. Penelitian ini menggunakan dataset publik Dreddit yang berisi teks pengguna Reddit berlabel stres dan non-stres. Penelitian ini bertujuan untuk membandingkan dua pendekatan deteksi stres berbasis teks, yaitu fine-tuning model transformer DeBERTa-v3 dan metode klasik TF-IDF LinearSVC. Kedua pendekatan dirancang sebagai sistem klasifikasi biner untuk mengenali teks stres dan non-stres secara otomatis. Tahapan penelitian meliputi praproses data, tokenisasi, pelatihan model, validasi, dan evaluasi menggunakan metrik Accuracy, Precision, Recall, F1-score, dan AUROC. Model DeBERTa-v3 dilatih menggunakan strategi fine-tuning berbasis representasi kontekstual dengan mekanisme self-attention, sedangkan TF-IDF LinearSVC mengandalkan pembobotan n-gram secara statistik. Hasil pengujian menunjukkan bahwa DeBERTa-v3 memberikan performa terbaik dengan Accuracy 0,830, Precision 0,802, Recall 0,889, F1-score 0,843, dan AUROC 0,918. Sementara itu, TF-IDF LinearSVC menghasilkan Accuracy 0,732, Precision 0,722, Recall 0,783, F1-score 0,751, dan AUROC 0,817. Eksperimen dilakukan dengan konfigurasi pelatihan, pembagian data, dan evaluasi yang konsisten sehingga perbandingan kedua model tetap adil. Analisis confusion matrix memperlihatkan bahwa DeBERTa-v3 memiliki false positive dan false negative yang lebih rendah, sehingga lebih efektif dalam mengenali ekspresi stres implisit dibandingkan pendekatan klasik. Penelitian ini menegaskan keunggulan model transformer dalam memahami konteks emosional dan semantik teks. Temuan ini membuka peluang penerapan model pada sistem pemantauan kesehatan mental berbasis media sosial secara real-time untuk deteksi dini kondisi psikologis masyarakat.

**Kata Kunci:** DeBERTa-v3; TF-IDF; LinearSVC; Deteksi Stres; NLP

## Abstract

Mental health has become an important issue in the digital era, as psychological expressions are increasingly reflected through social media posts such as Reddit. This study uses the publicly available Dreddit dataset containing Reddit user texts labeled with stress categories. The objective is to compare two text-based stress detection approaches: fine-tuning the transformer model DeBERTa-v3 and the classical TF-IDF LinearSVC method. Both approaches are implemented as binary classification systems to automatically distinguish stress and non-stress texts. The research workflow includes data preprocessing, tokenization, model training, validation, and evaluation using Accuracy, Precision, Recall, F1-score, and AUROC metrics. DeBERTa-v3 is fine-tuned using contextual representations with a self-attention mechanism, while TF-IDF LinearSVC relies on statistical n-gram weighting. Experimental results show that DeBERTa-v3 achieves superior performance with an Accuracy of 0.830, Precision of 0.802, Recall of 0.889, F1-score of 0.843, and AUROC of 0.918. Meanwhile, TF-IDF LinearSVC obtains an Accuracy of 0.732, Precision of 0.722, Recall of 0.783, F1-score of 0.751, and AUROC of 0.817. The experiments were conducted with consistent training configurations, data splits, and evaluation procedures to ensure a fair comparison. The confusion matrix analysis indicates that DeBERTa-v3 produces fewer false positives and false negatives, demonstrating stronger capability in recognizing implicit stress expressions. These findings highlight the advantages of transformer-based models in capturing emotional and semantic context and indicate the potential for real-time deployment in social-media-based mental health monitoring systems.

**Keywords:** DeBERTa-v3; TF-IDF; LinearSVC; Stress Detection; Natural Language Processing

## 1. PENDAHULUAN

Kesehatan mental merupakan aspek penting dalam kehidupan manusia yang kini mendapat perhatian serius di berbagai negara. Laporan World Health Organization (WHO) tahun 2023 menyebutkan bahwa satu dari delapan orang di dunia hidup dengan gangguan mental seperti stres, kecemasan, dan depresi, dengan total beban ekonomi global mencapai lebih dari satu triliun dolar setiap tahun [1], [2], [3]. Di era digital, ekspresi psikologis individu kerap terekam melalui media sosial, termasuk Reddit, yang memiliki jutaan pengguna aktif setiap harinya [4]. Reddit menjadi ruang bagi banyak pengguna untuk mencurahkan perasaan, berdiskusi, dan berbagi pengalaman, sehingga teks yang dihasilkan mengandung sinyal linguistik yang dapat merefleksikan kondisi mental seseorang [5], [6], [7]. Laporan Gallup Global Emotions menunjukkan bahwa tingkat stres global mencapai rekor tertinggi dalam satu dekade, memperkuat urgensi pengembangan metode deteksi stres berbasis data untuk konteks digital modern [8].

Deteksi stres berbasis teks bertujuan untuk mengidentifikasi indikator linguistik yang menunjukkan tekanan psikologis dari unggahan pengguna. Namun, pendekatan manual melalui analisis psikologis masih memerlukan waktu dan sumber daya yang besar, serta sulit diterapkan pada volume data yang sangat tinggi [9], [10]. Oleh karena itu, pendekatan pembelajaran mesin banyak digunakan untuk menganalisis pola bahasa dan konteks kalimat secara otomatis [11], [12]. Salah satu dataset publik yang paling sering digunakan adalah Dreddit, yaitu kumpulan unggahan Reddit yang berlabel stres dan non-stres, yang menyediakan dasar bagi pengujian berbagai metode klasifikasi teks dalam domain kesehatan mental [13].

Berbagai penelitian sebelumnya telah berupaya melakukan deteksi stres dan gangguan psikologis berbasis teks menggunakan pendekatan Natural Language Processing (NLP) dan pembelajaran mesin [14], [15]. Metode berbasis fitur klasik seperti TF-IDF dan Bag-of-Words masih banyak digunakan dalam tahap awal pengembangan model karena kemudahannya dalam menangani data teks, namun performanya sering kali kurang memuaskan saat dihadapkan pada kalimat yang bersifat ambigu atau mengandung makna kontekstual [5], [16]. Pendekatan lain mencoba memadukan fitur leksikal dengan representasi distribusional seperti FastText untuk memperkaya konteks semantik, namun hasilnya tetap menunjukkan keterbatasan dalam menangkap emosi implisit dari pengguna [17]. Kajian sistematis juga menunjukkan bahwa mayoritas penelitian NLP dalam bidang kesehatan mental masih mengandalkan data media sosial dan catatan klinis dengan teknik pra-pemrosesan dasar seperti n-gram, lemmatization, dan part-of-speech tagging, serta algoritma seperti SVM dan Logistic Regression yang memiliki interpretabilitas baik namun kurang optimal dalam pemodelan konteks panjang [18]. Pendekatan berbasis analisis topik dan sentimen turut digunakan untuk menelusuri tema-tema gangguan mental di platform daring seperti Reddit, tetapi hasilnya cenderung deskriptif dan belum cukup efektif untuk klasifikasi otomatis [19]. Oleh karena itu, muncul kebutuhan akan model yang tidak hanya memahami kata secara leksikal, tetapi juga mampu mengenali konteks emosional dan relasi semantik antar kata dalam mendeteksi stres berbasis teks.

Pendekatan modern berbasis transformer kemudian mulai digunakan untuk mengatasi keterbatasan tersebut. Model berarsitektur BERT dan turunannya seperti RoBERTa serta MentalRoBERTa terbukti memberikan hasil yang lebih baik dibandingkan pendekatan berbasis fitur tradisional karena kemampuannya memahami makna kontekstual dan emosi secara mendalam [18], [20], [21], [22]. Penelitian lain mengusulkan model KC-Net yang menggabungkan knowledge-aware dan contrastive learning, dan terbukti lebih stabil serta mampu mempelajari hubungan semantik antar kalimat dengan lebih baik [9]. Beberapa studi juga mengadopsi pendekatan multi-task learning untuk mendeteksi stres dan depresi secara bersamaan, yang dinilai memberikan hasil yang lebih seimbang dan generalisasi yang lebih baik terhadap data baru [23]. Selain itu, metode emotion-infused fine-tuning pada model transformer berhasil meningkatkan sensitivitas model terhadap konteks emosional pengguna dibandingkan model transformer dasar [13]. Adapun penelitian dengan model IndoBERT menunjukkan hasil yang baik dalam bahasa Indonesia, namun masih terbatas pada cakupan konteks linguistik tertentu [24].

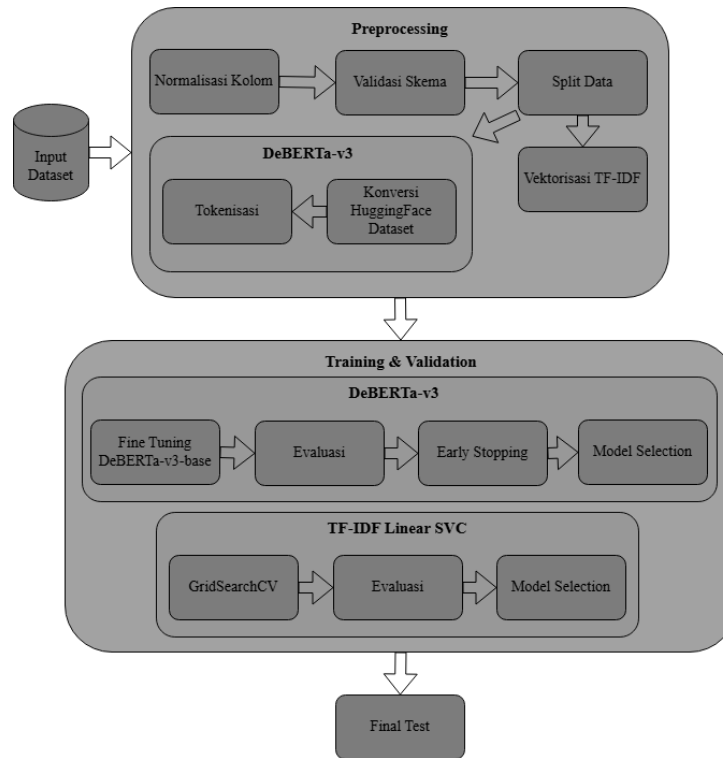
Meskipun begitu, masih terdapat kesenjangan penelitian yang cukup jelas. Sebagian besar studi terdahulu belum melakukan perbandingan langsung antara model transformer terbaru DeBERTa-v3 dan TF-IDF LinearSVC dengan protokol evaluasi yang konsisten pada Dreddit [25], [26]. Selain itu, sebagian penelitian masih berfokus pada depresi atau kecemasan, bukan pada deteksi stres secara spesifik [27]. Faktor lain seperti pelaporan metrik per kelas dan analisis kesalahan (confusion matrix) juga sering diabaikan, padahal informasi tersebut penting untuk memahami sumber keunggulan atau kelemahan tiap model [28]. Dengan demikian, penelitian lanjutan diperlukan untuk menilai secara objektif keunggulan representasi kontekstual modern dibandingkan pendekatan berbasis fitur tradisional pada data Reddit. Pendahuluan ini juga menegaskan kontribusi penelitian berupa evaluasi komparatif yang lebih menyeluruh, termasuk analisis kesalahan dan pelaporan metrik per kelas, yang belum banyak dilakukan pada studi sebelumnya.

Secara umum, penelitian-penelitian terdahulu menegaskan bahwa pendekatan berbasis transformer lebih unggul dibandingkan metode klasik, tetapi belum banyak studi yang membandingkan secara langsung model transformer encoder modern seperti DeBERTa-v3 dengan pendekatan klasik TF-IDF LinearSVC pada dataset Dreddit. Oleh karena itu, penelitian ini bertujuan untuk menganalisis deteksi stres berbasis teks dengan membandingkan kinerja model transformer encoder modern melalui pendekatan fine-tuning DeBERTa-v3 dan metode klasik TF-IDF LinearSVC pada dataset Dreddit. Perbandingan ini dilakukan untuk menilai efektivitas representasi kontekstual modern dalam meningkatkan akurasi deteksi stres dibandingkan pendekatan berbasis fitur tradisional. Evaluasi dilakukan menggunakan metrik Macro-F1, Accuracy, dan AUROC untuk menilai keseimbangan kinerja antar kelas. Hasil yang diperoleh diharapkan dapat memberikan kontribusi teoretis dalam memperkuat bukti empiris efektivitas model transformer pada analisis kesehatan mental berbasis teks, serta manfaat praktis berupa pendekatan analitis yang dapat membantu peneliti dan lembaga terkait dalam memantau sinyal stres di platform daring secara lebih cepat, akurat, dan replikabel.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Penelitian ini membandingkan dua pendekatan untuk deteksi stres berbasis teks pada Dreddit yaitu fine-tuning DeBERTa-v3 dan TF-IDF dengan LinearSVC. Setiap proses memiliki metode – metode dan pendekatan yang lebih mendalam untuk memenuhi kebutuhan proses berikutnya. Tahapan eksperimen pada penelitian ini divisualisasikan melalui flowchart pada Gambar 1 untuk memperjelas urutan proses, mulai dari input dataset, preprocessing, training dan validasi, hingga evaluasi akhir.



**Gambar 1.** Tahapan Penelitian

## 2.1 Input Dataset

Penelitian ini menggunakan korpus Dreddit yang berisi postingan Reddit dari beragam domain misalnya Anxiety, PTSD, Homeless, dsb [13]. Pada penelitian ini, hanya satu fitur dan satu label yang digunakan yaitu teks posting, dan label biner dengan nilai 0 = non-stres dan 1 = stres.

## 2.2 Preprocessing

### 2.2.1 Normalisasi Kolom

Pada tahap awal seluruh nama kolom pada Dataset Dreddit dinormalisasi menjadi huruf kecil untuk menyeragamkan penamaan dan mencegah error yang bersifat case sensitive. Setelah itu dilakukan pemetaan eksplisit satu fitur teks yaitu kolom “text” dan satu target yaitu kolom “label”. Penelitian ini tidak menerapkan pembersihan linguistik agresif seperti stopword removal atau stemming karena pada cabang DeBERTa tokenizer subword sudah dirancang untuk menangani variasi ejaan dan tanda baca tanpa kehilangan informasi semantik penting sedangkan pada cabang TF IDF penghapusan stopword atau stemming berisiko menghilangkan sinyal n gram seperti negasi dan penguat sehingga dibiarkan apa adanya.

### 2.2.2 Validasi Skema

Setelah normalisasi dan pemilihan kolom, dilakukan validasi skema untuk memastikan kesesuaian dengan rancangan eksperimen. Validasi ini memeriksa bahwa dataset hanya memuat tepat satu fitur masukan berupa teks dengan nama “text” dan satu target keluaran dengan nama “label”, sekaligus menampilkan ringkasan jumlah baris, serta memastikan tidak ada kolom lain yang ikut terbawa ke pipeline. Hasil validasi dicatat ke berkas ringkasan sehingga dapat dijadikan bukti reproduktibilitas dan audit kualitas data.

### 2.2.3 Split Data

Dataset pelatihan dipecah menjadi subset pelatihan dan validasi dengan proporsi 80:20 menggunakan stratified split berdasarkan kolom “label”. Strategi ini menjaga proporsi kelas stres/non-stres tetap seimbang di kedua subset, sehingga evaluasi selama pelatihan mencerminkan kinerja model yang tidak bias terhadap kelas mayoritas. Untuk mendukung replikasi pemecahan data dikendalikan oleh random seed tetap, sementara himpunan uji disimpan terpisah dan tidak disentuh selama pelatihan.

## 2.3 DeBERTa-v3

Pemilihan DeBERTa-v3 pada penelitian ini didasarkan pada hasil studi sebelumnya yang menunjukkan bahwa arsitektur disentangled attention dan enhanced mask decoder miliknya memberikan performa lebih stabil dan akurat dibandingkan RoBERTa dan MentalRoBERTa dalam tugas klasifikasi teks emosional.

### 2.3.1 Konversi HuggingFace Dataset

Agar data dapat diproses langsung oleh Trainer dari pustaka Transformers, setiap pandas DataFrame (train, validation, dan test) dikonversi ke objek dataset dan dihimpun ke dalam sebuah DatasetDict. Konversi ini menyederhanakan alur komputasi, memungkinkan pemetaan fungsi tokenisasi secara vektorisasi, serta memastikan atribut yang diperlukan (misalnya “text” dan “label”) tersedia dalam format yang konsisten untuk seluruh tahapan.

### 2.3.2 Tokenisasi

Teks pada setiap subset kemudian ditokenisasi menggunakan AutoTokenizer yang sesuai dengan model dasar DeBERTa-v3. Proses tokenisasi menerapkan truncation dan padding dengan max\_length 256 untuk menyeimbangkan cakupan konteks dan efisiensi komputasi. Keputusan untuk tidak melakukan tokenisasi manual atau normalisasi leksikal tambahan didasarkan pada kemampuan tokenizer subword dalam menangani ragam ejaan, tanda baca, dan variasi bentuk kata, sehingga informasi penting tetap terjaga menjelang tahap pelatihan.

## 2.4 TF-IDF LinearSVC

### 2.4.1 Vektorisasi TF-IDF

Representasi teks dibangun dengan Term Frequency–Inverse Document Frequency untuk menimbang istilah berdasarkan kemunculan di dokumen dan kelangkaan pada korpus. Proses ini mencakup teks dinormalisasi menjadi huruf kecil serta aksentuasi pada huruf dihilangkan dan tokenisasi berbasis pola regex sehingga unit analisis berupa kata beserta n gram kata sesuai ruang penelusuran yang dipertimbangkan yaitu unigram bigram dan trigram. Penyusunan fitur menggunakan skala TF sublinear agar pengulangan kata yang berlebihan tidak mendominasi bobot. Pemfilteran dilakukan melalui batas minimum kemunculan dan batas maksimum proporsi dokumen serta dapat diberi batas jumlah fitur bila diperlukan.

## 2.5 Training & Validation DeBERTa-v3

### 2.5.1 Fine Tuning

Model DeBERTa-v3 dilatih untuk klasifikasi biner stres dan non-stres menggunakan Trainer dari Transformers. Evaluasi dilakukan pada data validasi setiap akhir epoch sehingga kemajuan model dapat dipantau dengan konsisten. Model terbaik ditentukan berdasarkan nilai Macro-F1 pada validasi. Prediksi untuk pelaporan utama diambil dengan argmax tanpa penalaan ambang. Parameter pelatihan seperti optimizer, ukuran batch, dan panjang token dapat dilihat pada Tabel 1. Tabel 1 berikut menyajikan konfigurasi hyperparameter yang digunakan pada proses fine-tuning DeBERTa-v3.

**Tabel 1.** Konfigurasi Fine Tuning DeBERTa-v3

Parameter	Nilai	Keterangan
Model dasar	DeBERTa-v3	Model pre-trained dari Microsoft untuk klasifikasi teks
Panjang token	256	Panjang maksimum input setelah tokenisasi
Optimizer	AdamW	Optimizer adaptif dengan weight decay
Learning rate	2e-5	Laju pembelajaran awal yang stabil untuk fine-tuning
Batch size	8	Jumlah data per iterasi pelatihan
Epoch	5	Jumlah iterasi penuh pada data pelatihan
Seed	42	Nilai acak agar hasil dapat direplikasi
Warmup ratio	0.1	Rasio langkah awal tanpa penurunan learning rate
Weight decay	0.01	Nilai regularisasi untuk menjaga stabilitas bobot
Metrik evaluasi	Accuracy, Macro-F1, AUROC	Digunakan untuk pemilihan model terbaik
Strategi prediksi	Argmax	Label akhir diambil dari probabilitas tertinggi

### 2.5.2 Evaluasi

Selama pelatihan sistem menghitung Accuracy dan Macro-F1 dari prediksi argmax pada data validasi. Nilai AUROC turut dihitung dengan memanfaatkan probabilitas kelas positif dari fungsi softmax agar kemampuan pemeringkatan tetap tercatat. Setelah pelatihan selesai metrik yang sama dihitung kembali pada himpunan validasi dan himpunan uji untuk pelaporan akhir. Laporan per kelas yang berisi precision recall dan F1 dihasilkan untuk menggambarkan kinerja pada masing-masing kelas.

### 2.5.3 Early Stopping

Mekanisme early stopping dengan dua epoch digunakan untuk mencegah overfitting[29]. Pelatihan dihentikan ketika tidak ada peningkatan pada metrik validasi dalam beberapa epoch terakhir sehingga waktu komputasi lebih efisien dan model berhenti pada titik generalisasi yang baik.

### 2.5.4 Model Selection

Pemilihan model terbaik dilakukan berdasarkan nilai Macro-F1 pada data validasi. Selama proses pelatihan, checkpoint disimpan pada akhir setiap epoch dan hanya checkpoint dengan Macro-F1 tertinggi yang dipertahankan sebagai kandidat

terbaik. Setelah selesai, bobot model dari checkpoint terbaik tersebut disalin sebagai best model dan digunakan untuk seluruh evaluasi lanjutan termasuk penyetelan ambang pada validasi serta pengukuran akhir di data uji.

## 2.6 Training & Validation TF-IDF LinearSVC

### 2.6.1 GridSearchCV

Pendekatan klasik dibangun sebagai pipeline vektorisasi TF-IDF diikuti pengklasifikasi LinearSVC. Ruang penelusuran mencakup pilihan stopwords, pengaturan rentang n gram, pengaturan ambang kemunculan minimum, dan proporsi maksimum batas jumlah fitur, serta penggunaan skala TF sublinear. Pada sisi pengklasifikasi dievaluasi beberapa nilai parameter C dan opsi pembobotan kelas. Penelusuran dilakukan dengan validasi silang lima lipatan menggunakan skor Macro-f1 sebagai kriteria dan pengaturan refit diaktifkan agar pipeline terbaik dilatih ulang pada data pelatihan penuh setelah penelusuran selesai.

### 2.6.2 Evaluasi

Pipeline terbaik dari hasil penelusuran dievaluasi pada himpunan validasi dan himpunan uji. Nilai yang dilaporkan meliputi Accuracy dan Macro-F1 serta AUROC yang dihitung dari margin decision function agar kemampuan pemeringkatan tetap tercakup. Selain itu dihasilkan laporan per kelas yang memuat precision recall dan F1 untuk menggambarkan distribusi kinerja antar kelas serta confusion matrix untuk memberikan gambaran kesalahan klasifikasi.

### 2.6.3 Model Selection

Pemilihan model pada pendekatan sepenuhnya didasarkan pada skor Macro-f1 dari validasi silang. Pipeline dengan nilai tertinggi menjadi estimator akhir dan digunakan untuk prediksi pada data uji. Evaluasi berupa metrik laporan dan confusion matrix disimpan agar mudah direplikasi.

## 2.7 Final Test

Setelah batas keputusan ditetapkan, pada tahap akhir kedua model dievaluasi pada himpunan uji yang tidak digunakan selama pelatihan. Untuk DeBERTa prediksi kelas diperoleh dari argmax atas logits tanpa penalaan ambang. Untuk TF-IDF LinearSVC label dihasilkan langsung oleh pengklasifikasi tanpa penalaan ambang. Nilai yang dilaporkan mencakup Accuracy Macro-F1 dan AUROC. Laporan per kelas yang berisi precision recall dan F1 serta confusion matrix untuk menggambarkan pola kesalahan klasifikasi. Evaluasi dilakukan satu kali pada susunan data yang sama sehingga hasilnya dapat direplikasi.

## 3. HASIL DAN PEMBAHASAN

Penelitian ini membandingkan dua pendekatan deteksi stres berbasis teks pada Dreddit yaitu fine-tuning DeBERTa-v3 dan TF-IDF LinearSVC. Metrik utama yang dilaporkan adalah Macro-F1 yang menilai keseimbangan kinerja antar kelas, sedangkan Accuracy dan AUROC disertakan sebagai metrik pelengkap. Prediksi pada DeBERTa diambil dengan argmax atas logits tanpa penalaan ambang, sedangkan TF-IDF LinearSVC menghasilkan label langsung dari pengklasifikasi.

### 3.1 Dataset

Dataset yang digunakan berasal dari Dreddit, yaitu korpus publik yang dikembangkan oleh Turcan dan McKeown untuk studi deteksi stres berbasis teks. Dataset ini berisi unggahan Reddit dari berbagai komunitas seperti Anxiety, PTSD, Relationships, dan Homeless dengan setiap entri terdiri dari satu teks posting dan satu label biner (0 = non-stres, 1 = stres). Pada penelitian ini digunakan 2.838 data pelatihan dan 715 data pengujian. Distribusi label pada kedua subset relatif seimbang sehingga mendukung proses pembelajaran model. Tabel 2 menyajikan contoh data untuk memperlihatkan konteks awal korpus Dreddit yang digunakan.

Tabel 2. Dataset

Text	Label
October is Domestic Violence Awareness Month and I am a domestic violence survivor who is still struggling, even after over four years....	1
It was a big company so luckily I didn't have to see him all the time, but when I did, he again acted as though I didn't exist. I tried to talk to him and update him on the pregnancy ...	0
My mom then hit me with the newspaper and it shocked me that she would do this, she knows I don't like play hitting, smacking, striking, hitting or violence of any sort on my person	1

### 3.2 Preprocessing

Tahap preprocessing dilakukan untuk menyiapkan data dalam format yang seragam. Seluruh nama kolom dinormalisasi menjadi huruf kecil, kemudian dipetakan secara eksplisit menjadi satu fitur teks dan satu label keluaran. Penelitian ini tidak menerapkan stopword removal atau stemming, karena tokenizer subword pada DeBERTa mampu mempertahankan informasi semantik penting, dan pada TF-IDF penghapusan stopword justru dapat menghilangkan pola n-gram yang relevan. Validasi skema dilakukan untuk memastikan dataset hanya berisi kolom “text” dan “label” serta tidak terdapat nilai label di luar 0 dan 1.

Setelah skema dinyatakan valid, dataset pelatihan dipecah menggunakan stratified split dengan rasio 80% data pelatihan dan 20% data validasi untuk menjaga keseimbangan kelas pada kedua subset. Dataset pengujian dipisahkan sejak awal dan hanya digunakan pada tahap akhir evaluasi. Pada tahap pembentukan fitur, teks ditokenisasi menggunakan AutoTokenizer milik DeBERTa-v3. Tokenizer ini memecah teks ke dalam unit subword dan menerapkan padding serta truncation hingga panjang maksimum 256 token. Untuk memberikan gambaran proses tokenisasi, contoh hasil tokenisasi ditampilkan pada Tabel 3.

**Tabel 3.** Hasil Tokenisasi

Text	Tokens	Length
October is Domestic Violence Awareness Month and I am a domestic violence survivor who is still struggling, even after over four years....	[October, is, Domestic, Violence, Awareness, Month, and, I, am, a, domestic, violence, survivor, who, is, still, struggling, even, after, over, four, years, ...]	107

Tabel tersebut memperlihatkan bagaimana teks asli diubah menjadi token subword serta ID numerik yang sesuai dengan kamus model. Informasi ini membantu menjelaskan proses representasi input sebelum memasuki tahap pelatihan model.

### 3.3 Tahapan Penerapan Algoritma

Tahapan penerapan algoritma terdiri dari dua alur utama, yaitu fine-tuning DeBERTa-v3 dan pemodelan klasik dengan TF-IDF LinearSVC. Pada pendekatan DeBERTa-v3, dataset yang telah ditokenisasi dikonversi ke dalam DatasetDict sehingga kompatibel dengan modul Trainer. Model kemudian dilatih menggunakan konfigurasi hyperparameter yang mencakup learning rate, batch size, warmup ratio, serta early stopping untuk menghindari overfitting. Melalui mekanisme self-attention, model mempelajari hubungan semantik dan emosional antarkata sehingga mampu mengenali pola stres secara lebih mendalam.

Sementara itu, pada pendekatan TF-IDF LinearSVC, proses vektorisasi teks dilakukan menggunakan TfidfVectorizer dengan kombinasi pengaturan seperti n-gram, min\_df, max\_df, sublinear term frequency, serta opsi stopwords. Pemilihan konfigurasi terbaik dilakukan menggunakan GridSearchCV lima lipatan dengan metrik Macro-f1 sebagai acuan. Pipeline terbaik hasil penelusuran kemudian digunakan sebagai estimator akhir untuk melakukan prediksi pada data validasi dan pengujian. Kedua model yang diperoleh melalui tahapan ini selanjutnya dievaluasi pada bagian pembahasan hasil.

### 3.4 Hasil Set Uji

Tabel 4 berikut menyajikan ringkasan performa kedua model pada himpunan uji Dreddit menggunakan tiga metrik utama, yaitu Accuracy, Macro-F1, dan AUROC. Metrik ini digunakan untuk menilai ketelitian, keseimbangan kinerja antar kelas, dan kemampuan pemeringkatan model dalam membedakan teks stres dan non-stres.

**Tabel 4.** Kinerja pada Set Uji Dreddit

Model	Accuracy	Macro-F1	AUROC
DeBERTa-v3	0.830	0.828	0.918
TF-IDF LinearSVC	0.732	0.731	0.817

DeBERTa-v3 menunjukkan performa yang lebih unggul dibandingkan TF-IDF LinearSVC pada seluruh metrik evaluasi. Nilai Macro-F1 yang lebih tinggi mengindikasikan bahwa model berbasis representasi kontekstual mampu menangkap hubungan semantik antar kata secara lebih efektif, termasuk ekspresi stres yang bersifat implisit. Sebaliknya, TF-IDF LinearSVC bergantung pada kemunculan literal n-gram, sehingga kurang stabil ketika pola bahasa mengandung makna ambigu atau tidak langsung. Temuan ini menegaskan bahwa pendekatan kontekstual lebih efektif dalam mendukung deteksi stres berbasis teks pada media sosial.

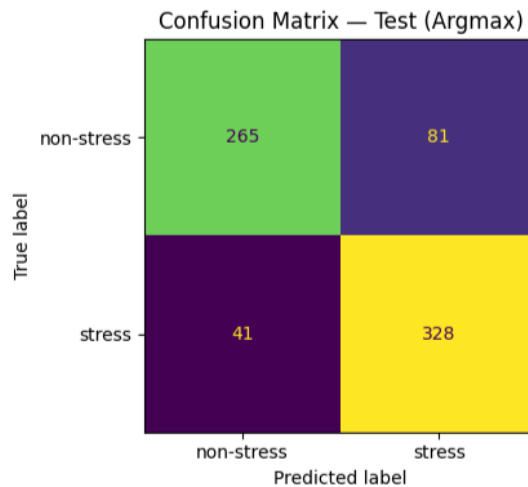
### 3.5 Analisis per kelas & Confusion Matrix

Tabel 5 berikut menampilkan metrik evaluasi per kelas untuk kategori stres dan non-stres, sehingga memberikan gambaran lebih detail mengenai seberapa baik kedua model mengenali masing-masing kelas. Evaluasi per kelas penting karena distribusi kesalahan tidak selalu simetris antara dua kelas.

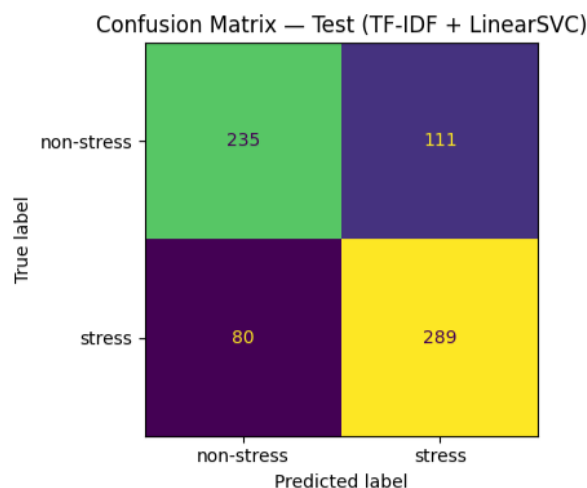
**Tabel 5.** Kinerja per kelas pada Set Uji Dreddit

Model	Label	Presisi	Recall	F1
DeBERTa-v3	Stres	0.802	0.889	0.843
	Non-stres	0.866	0.766	0.813
TF-IDF LinearSVC	Stres	0.722	0.783	0.751
	Non-stres	0.746	0.679	0.711

Model DeBERTa-v3 menunjukkan recall yang jauh lebih tinggi pada kelas stres, yaitu 0,889 dibandingkan 0,783 pada TF-IDF LinearSVC. Nilai recall yang tinggi mengindikasikan bahwa model lebih sensitif dalam menangkap ekspresi stres sehingga lebih sedikit kasus stres yang terlewat (false negative). Precision DeBERTa-v3 pada kelas stres juga lebih tinggi (0,802 vs 0,722), menunjukkan bahwa model kontekstual lebih jarang salah dalam mengklasifikasikan teks non-stres sebagai stres, sehingga jumlah false positive lebih rendah. Untuk kelas non-stres, performa kedua model relatif seimbang, meskipun DeBERTa-v3 tetap unggul pada F1, yang menunjukkan kemampuan yang lebih stabil dalam mempertahankan keseimbangan precision dan recall. Gambar 2 dan Gambar 3 menampilkan confusion matrix hasil klasifikasi model DeBERTa-v3 dan TF-IDF LinearSVC.



**Gambar 2.** Confusion Matrix DeBERTa-v3



**Gambar 3.** Confusion Matrix TF-IDF LinearSVC

Berdasarkan Gambar 2 dan Gambar 3, Pengujian dilakukan menggunakan data uji untuk menilai kinerja model dan evaluasi performa menggunakan confusion matrix. Confusion matrix banyak digunakan dalam analisis data untuk mengukur tingkat akurasi dan menilai kinerja model klasifikasi secara keseluruhan [30]. Dapat dilihat jumlah prediksi

benar dan salah untuk masing-masing kelas. Nilai-nilai ini menjadi dasar perhitungan metrik evaluasi menggunakan rumus dasar confusion matrix berikut:

$$\text{Presisi} = \frac{TP}{(TP+FP)} \quad (1)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (2)$$

$$F1 - \text{score} = 2 \times \frac{(\text{Presisi} \times \text{Recall})}{(\text{Presisi} + \text{Recall})} \quad (3)$$

Berdasarkan Gambar 2 (DeBERTa-v3), diperoleh nilai TP = 328, FP = 81, FN = 41, dan TN = 265. Maka:

$$\text{Presisi DeBERTa - v3} = \frac{328}{(328+81)} = 0.802 \quad (4)$$

$$\text{Recall DeBERTa - v3} = \frac{328}{(328+41)} = 0.889 \quad (5)$$

$$F1 \text{ DeBERTa - v3} = 2 \times \frac{(0.802 \times 0.889)}{(0.802 + 0.889)} = 0.843 \quad (6)$$

Berdasarkan Gambar 3 (TF-IDF LinearSVC), diperoleh nilai TP = 289, FP = 111, FN = 80, dan TN = 235. Maka:

$$\text{Presisi DeBERTa - v3} = \frac{289}{(289+111)} = 0.722 \quad (7)$$

$$\text{Recall DeBERTa - v3} = \frac{289}{(289+80)} = 0.783 \quad (8)$$

$$F1 \text{ DeBERTa - v3} = 2 \times \frac{(0.722 \times 0.783)}{(0.722 + 0.783)} = 0.751 \quad (9)$$

Hasil perhitungan tersebut konsisten dengan nilai metrik yang diperoleh pada pengujian. Model DeBERTa menghasilkan false negative lebih sedikit 41 dibanding 80 dan false positive lebih sedikit 81 dibanding 111, yang menjelaskan peningkatan recall dan precision sekaligus. Dengan Macro-F1 sebesar 0,828 dan AUROC 0,918, model DeBERTa memperlihatkan keseimbangan terbaik antara ketelitian dan ketercakupannya prediksi. Jika ditinjau lebih lanjut dari confusion matrix, kesalahan klasifikasi yang paling sering terjadi adalah pada teks non-stres yang diklasifikasikan sebagai stres. Pola ini mengindikasikan bahwa sebagian besar kesalahan berasal dari kalimat yang mengandung kata-kata bermakna ambigu seperti “tired”, “anxious”, atau “overwhelmed” yang sering muncul baik pada konteks normal maupun stres. Namun, jumlah false negative yang relatif kecil memperlihatkan bahwa model lebih sensitif terhadap tanda-tanda stres dibandingkan kehilangan kasus stres yang sebenarnya. Hal ini menjadi penting dalam konteks deteksi dini, karena lebih baik mendeteksi sedikit berlebih dibandingkan gagal mengenali individu yang mengalami tekanan emosional.

### 3.6 Hasil Visualisasi Word Cloud

Visualisasi word cloud menampilkan kata-kata yang paling sering muncul dalam data teks Dreddit berdasarkan dua kategori utama, yaitu stres dan non-stres. Word cloud merupakan representasi visual dari frekuensi kemunculan istilah dalam korpus, di mana ukuran kata yang lebih besar menandakan tingkat kemunculan yang lebih tinggi. Visualisasi ini berperan penting dalam memahami pola linguistik yang sering digunakan oleh pengguna Reddit dalam mengekspresikan kondisi emosionalnya. Dengan cara ini, word cloud menjadi alat eksploratif yang efisien untuk mengidentifikasi topik dominan, emosi yang tersirat, serta perbedaan penggunaan bahasa antara kelompok yang mengalami stres dan yang tidak. Hasil ini juga membantu peneliti memperoleh pemahaman awal sebelum dilakukan analisis yang lebih mendalam terhadap konteks semantik dan sentimen dalam teks. Secara keseluruhan, visualisasi ini memperkuat interpretasi model dalam mengenali kata-kata yang menjadi indikator stres maupun ekspresi netral di media sosial Reddit. Gambar 4 menyajikan visualisasi word cloud dari hasil pemodelan DeBERTa-v3.

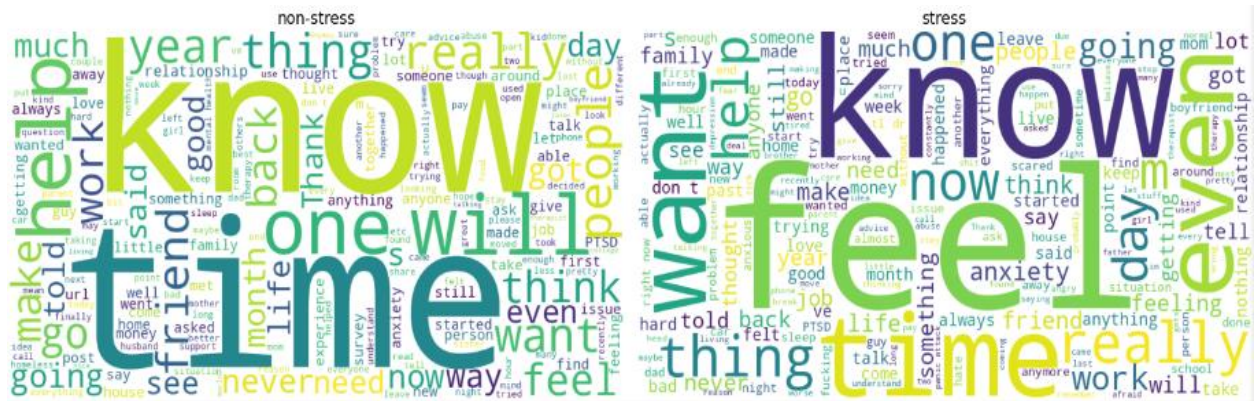


**Gambar 4.** Word Cloud DeBERTa-v3

Gambar 4 memperlihatkan hasil word cloud dari model DeBERTa-v3 pada kategori non-stres dan stres. Pada kelompok non-stres, kata-kata seperti know, time, help, people, work, dan friend muncul dengan ukuran besar. Hal ini menunjukkan bahwa teks dari pengguna dengan kategori non-stres cenderung menampilkan bahasa yang bersifat sosial,



informatif, dan netral. Pengguna lebih banyak berbagi cerita atau berdiskusi mengenai aktivitas sehari-hari, hubungan pertemanan, serta pengalaman hidup tanpa ekspresi emosional yang intens. Sementara itu, pada kategori stres, kata-kata dominan seperti feel, need, really, want, bad, anxiety, dan work menunjukkan adanya peningkatan ekspresi emosional dan tekanan psikologis. Kemunculan kata feel yang sangat menonjol mengindikasikan bahwa pengguna cenderung menulis tentang perasaan pribadi dan pengalaman emosional secara langsung. Selain itu, kata need dan help sering muncul bersamaan dengan konteks permintaan dukungan atau curahan hati. Pola ini menunjukkan bahwa model DeBERTa-v3 mampu menangkap nuansa emosional yang lebih dalam melalui representasi kontekstualnya, sehingga membantu mendeteksi ekspresi stres implisit yang tidak selalu tampak dari kata-kata eksplisit seperti “stres” atau “anxiety”. Gambar 5 menyajikan visualisasi word cloud dari hasil pemodelan TF-IDF LinearSVC.



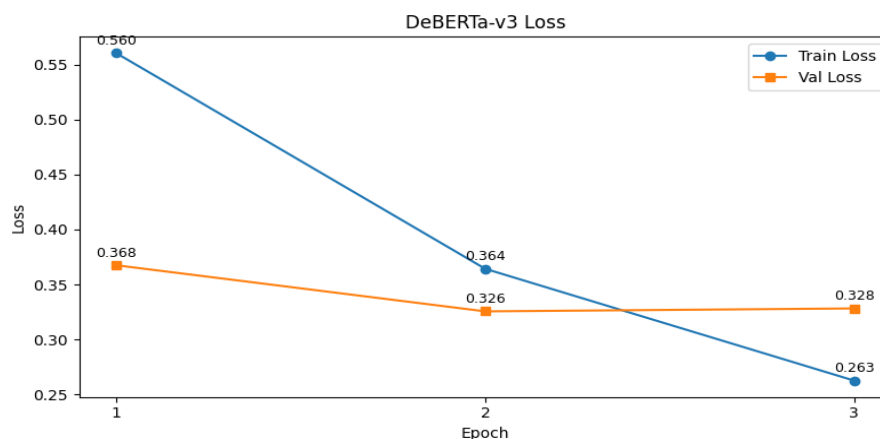
Gambar 5. Word Cloud TF-IDF LinearSVC

Gambar 5 menunjukkan hasil visualisasi dari model TF-IDF LinearSVC yang menggambarkan distribusi kata pada kategori non-stres dan stres. Secara umum, hasil word cloud pada model ini menampilkan kemiripan kata dominan dengan model DeBERTa-v3, namun dengan distribusi yang lebih sederhana. Pada kelompok non-stres, kata seperti know, time, work, people, life, dan good mendominasi, mencerminkan gaya bahasa yang lebih faktual dan deskriptif tanpa muatan emosional yang kuat. Sementara pada kelompok stres, kata seperti feel, want, really, now, help, dan anxiety lebih menonjol. Meskipun pola umum serupa, model TF-IDF LinearSVC cenderung menangkap kata berdasarkan frekuensi literal tanpa mempertimbangkan hubungan antar konteks. Hal ini menyebabkan model klasik ini kurang mampu membedakan perbedaan makna yang halus antara kata yang muncul pada konteks emosional dan konteks netral. Perbandingan ini memperlihatkan bahwa pendekatan berbasis vektor kontekstual seperti DeBERTa-v3 lebih unggul dalam mengenali emosi implisit, sedangkan TF-IDF LinearSVC hanya mampu memetakan perbedaan berdasarkan frekuensi kemunculan kata semata.

Secara keseluruhan, visualisasi word cloud dari kedua model ini memberikan pemahaman tambahan mengenai karakteristik linguistik yang membedakan teks stres dan non-stres. Word cloud menjadi representasi visual yang efektif untuk melihat pola umum pada korpus data, mendukung analisis kuantitatif yang dilakukan melalui metrik evaluasi seperti precision, recall, dan F1-score. Temuan ini juga memperkuat hasil analisis model yang menunjukkan bahwa DeBERTa-v3 memiliki sensitivitas lebih tinggi terhadap makna kontekstual dan emosi tersirat dibandingkan TF-IDF LinearSVC yang bekerja berdasarkan pendekatan statistik sederhana.

### 3.7 Analisis Validasi dan Pemilihan Model

Gambar 6 menyajikan kurva perubahan nilai loss selama proses pelatihan model DeBERTa-v3 pada data pelatihan dan validasi. Grafik ini digunakan untuk menilai stabilitas proses fine-tuning serta potensi terjadinya overfitting.



Gambar 6. Kurva Pelatihan dan Validasi Model DeBERTa-v3

Nilai train loss mengalami penurunan bertahap dari 0,560 menjadi 0,263, sedangkan validation loss turun dari 0,368 menjadi 0,326 sebelum stabil pada 0,328 di epoch ketiga. Pola penurunan kedua kurva yang relatif sejajar menunjukkan bahwa proses pembelajaran berlangsung stabil tanpa indikasi overfitting yang berarti. Konsistensi ini sejalan dengan hasil evaluasi kuantitatif, di mana nilai Macro-F1 pada validasi sebesar 0,871 hanya sedikit menurun menjadi 0,828 pada data uji. Nilai AUROC yang tetap tinggi pada kedua set data juga memperlihatkan bahwa kemampuan model dalam membedakan kelas stres dan non-stres tetap terjaga dengan baik.

Tabel 6 merangkum hasil evaluasi kedua pendekatan, yaitu DeBERTa-v3 dan TF-IDF LinearSVC, pada data validasi dan data uji untuk menilai performa serta generalisasi model pada susunan data yang berbeda.

**Tabel 6.** Hasil Validasi dan Uji DeBERTa-v3 dan TF-IDF LinearSVC

Model	Dataset	Accuracy	Macro-F1	AUROC
DeBERTa-v3	Validasi	0.871	0.871	0.944
	Tes	0.830	0.828	0.918
TF-IDF LinearSVC	Validasi	0.750	0.749	0.837
	Tes	0.733	0.731	0.818

Evaluasi performa menunjukkan Macro-F1 pada data validasi sebesar 0,871 dan sedikit menurun menjadi 0,828 pada data uji. Penurunan ini tergolong kecil sehingga model memiliki generalisasi yang baik. Nilai AUROC yang tetap tinggi 0,944 pada validasi dan 0,918 pada uji menunjukkan kemampuan model membedakan kelas stres dan non-stres secara konsisten.

Untuk TF-IDF LinearSVC, validasi lima lipatan menghasilkan akurasi 0,750 dan Macro-F1 0,749 pada validasi, lalu 0,733 dan 0,731 pada uji. Perbedaan hasil menunjukkan performa yang stabil di antara dua set data. Parameter terbaik diperoleh dari kombinasi rentang n-gram, min\_df, max\_df, penggunaan stopwords, dan pembobotan kelas yang memberikan keseimbangan antara bias dan varians. Selain performa yang lebih tinggi, model DeBERTa-v3 membutuhkan waktu pelatihan yang jauh lebih lama dan sumber daya komputasi yang lebih besar dibandingkan TF-IDF LinearSVC. Namun, peningkatan performa yang signifikan pada metrik F1 dan AUROC membenarkan investasi komputasi tersebut, terutama untuk aplikasi yang memerlukan deteksi stres secara akurat di dunia nyata. Sebaliknya, TF-IDF LinearSVC tetap relevan untuk skenario dengan keterbatasan sumber daya karena waktu latih yang cepat dan hasil yang masih kompetitif.

### 3.8 Perbandingan Hasil

Bagian ini membandingkan hasil penelitian dengan beberapa model terdahulu yang diuji pada dataset Dreddit. Tabel 7 menampilkan nilai presisi, recall, dan F1 dari masing-masing metode.

**Tabel 7.** Perbandingan Kinerja Model

Dataset	Model	Presisi	Recall	F1
Dreddit	MentalRoBERTa	0.821	0.818	0.819
	KC-Net+RoBERTa	0.827	0.826	0.827
	KC-Net	0.841	0.833	0.835
	Bag-of-Words	0.594	0.672	0.630
	TF-IDF	0.617	0.749	0.646
	DeBERTa-v3	<b>0.802</b>	<b>0.889</b>	<b>0.843</b>
	TF-IDF LinearSVC	<b>0.722</b>	<b>0.783</b>	<b>0.751</b>

Berdasarkan Tabel 7, model DeBERTa-v3 masih menunjukkan performa terbaik dengan nilai F1 sebesar 0,843, melampaui seluruh pendekatan lain termasuk model klasik dan transformer sebelumnya. Nilai recall yang tinggi, yaitu 0,889, menunjukkan kemampuan model ini dalam mengenali ekspresi stres yang tidak selalu menggunakan kata kunci eksplisit. Sementara itu, model berbasis fitur klasik seperti Bag-of-Words dan TF-IDF menghasilkan performa yang relatif lebih rendah, dengan F1 masing-masing sebesar 0,630 dan 0,646. Hasil ini menunjukkan bahwa metode berbasis frekuensi kata cenderung kurang mampu menangkap makna kontekstual dan relasi semantik antarkata. Meskipun demikian, TF-IDF LinearSVC masih memberikan hasil kompetitif dengan F1 sebesar 0,751, menunjukkan bahwa pendekatan n-gram yang divalidasi dengan pembobotan dan optimasi parameter tetap efektif pada data berukuran sedang. Jika dibandingkan dengan penelitian terdahulu, hasil DeBERTa-v3 berada dalam rentang performa model transformer modern seperti KC-Net+RoBERTa dan KC-Net, yang juga menonjol pada tugas deteksi stres berbasis teks. Berdasarkan hasil keseluruhan, perbedaan kinerja antara pendekatan kontekstual dan leksikal tidak hanya terlihat dari nilai metrik, tetapi juga pada cara kedua model memahami bahasa alami. Model DeBERTa-v3 yang dilatih melalui mekanisme self-attention mampu menangkap hubungan semantik antar kata meskipun posisinya berjauhan, sehingga lebih peka terhadap emosional yang kompleks. Sebaliknya, TF-IDF LinearSVC masih mengandalkan frekuensi kata dan urutan literal, yang membuatnya kesulitan mendeteksi makna implisit seperti ironi, sindiran, atau perasaan yang disamarkan. Penelitian ini memperkuat argumentasi bahwa pendekatan berbasis transformer tidak hanya memberikan peningkatan kuantitatif, tetapi juga

kualitatif dalam memahami konteks psikologis pengguna. Dengan demikian, DeBERTa-v3 terbukti lebih representatif untuk menganalisis fenomena sosial linguistik di platform daring yang bersifat informal dan penuh ekspresi emosional.

#### 4. KESIMPULAN

Penelitian ini membandingkan dua pendekatan deteksi stres berbasis teks pada dataset Dreddit, yaitu fine-tuning model DeBERTa-v3 dan pendekatan metode klasik TF-IDF LinearSVC. Berdasarkan hasil pengujian, model DeBERTa-v3 menunjukkan kinerja yang paling unggul dengan nilai Macro-F1 sebesar 0,828 dan AUROC 0,918, melampaui performa TF-IDF LinearSVC serta metode tradisional lainnya. Temuan ini juga menunjukkan bahwa DeBERTa-v3 mengungguli model-model terdahulu seperti MentalRoBERTa, KC-Net+RoBERTa, dan bahkan KC-Net, sehingga menegaskan kontribusi baru berupa capaian F1 tertinggi yaitu 0,843 di antara penelitian sebelumnya. Hasil ini menunjukkan bahwa representasi kontekstual pada model transformer mampu memahami hubungan semantik dan emosional antar kata, sehingga dapat mengenali ekspresi stres yang tidak selalu bersifat eksplisit. Selain peningkatan F1, model DeBERTa-v3 juga mencatat recall tertinggi yaitu 0,889, melebihi seluruh model terdahulu yang mengindikasikan sensitivitas lebih baik dalam mendeteksi stres implisit dibandingkan arsitektur transformer generasi sebelumnya. Sebaliknya, pendekatan berbasis frekuensi kata seperti TF-IDF atau Bag-of-Words masih menunjukkan keterbatasan dalam menangkap konteks dan makna implisit meskipun tetap memberikan hasil yang kompetitif pada data berukuran sedang. Selain itu, proses validasi dan pemilihan model menunjukkan stabilitas hasil pada data validasi dan uji, yang menandakan kemampuan generalisasi yang baik dari model DeBERTa-v3. Penggunaan confusion matrix turut memperkuat analisis dengan memperlihatkan bahwa DeBERTa-v3 memiliki jumlah false positive dan false negative yang lebih rendah dibanding TF-IDF LinearSVC. Dengan demikian, dapat disimpulkan bahwa model berbasis transformer modern memberikan hasil yang lebih akurat dan seimbang dalam tugas deteksi stres berbasis teks. Namun, penelitian ini memiliki beberapa keterbatasan, di antaranya jumlah data yang terbatas dan penggunaan dataset tunggal, sehingga generalisasi model perlu diuji lebih lanjut pada domain lain. Penelitian mendatang disarankan untuk memperluas dataset, menguji model pada bahasa dan platform berbeda, serta mengeksplorasi integrasi fitur multimodal seperti gambar atau metadata pengguna guna meningkatkan ketepatan dan sensitivitas model dalam mendeteksi stres secara lebih komprehensif.

#### REFERENCES

- [1] WHO, "Transforming mental health for all EXECUTIVE SUMMARY," *World Ment. Health Rep.*, 2022.
- [2] C. Su, Z. Xu, J. Pathak, dan F. Wang, "Deep learning in mental health outcome research : a scoping review," *Transl. Psychiatry*, 2020, doi: 10.1038/s41398-020-0780-3.
- [3] C. Sunkel, "The World Mental Health Report : transforming mental health for all," hlm. 391–392, 2001, doi: 10.1002/wps.21018.
- [4] T. Nijhawan, G. Attigeri, dan T. Ananthkrishna, "Stress detection using natural language processing and machine learning over social interactions," *J. Big Data*, vol. 9, no. 1, hlm. 33, Des 2022, doi: 10.1186/s40537-022-00575-6.
- [5] S. Inamdar, R. Chapekar, S. Gite, dan B. Pradhan, "Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing," *Hum.-Centric Intell. Syst.*, vol. 3, no. 2, hlm. 80–91, 2023, doi: 10.1007/s44230-023-00020-8.
- [6] S. M. Joseph, S. Citraro, V. Morini, G. Rossetti, dan M. Stella1, "Cognitive network neighbourhoods quantify feelings expressed in suicide notes and Reddit mental health communities," *ScienceDirect*, 2023.
- [7] D. F. Risa, F. Pradana, F. A. Bachtiar, dan U. Brawijaya, "IMPLEMENTASI METODE NAÏVE BAYES UNTUK MENDETEKSI STRES SISWA IMPLEMENTATION OF THE NAÏVE BAYES METHOD FOR DETECTING STUDENT," vol. 8, no. 6, 2021, doi: 10.25126/jtiik.202184372.
- [8] G. Inc, "State of the World's Emotional Health Report," Gallup.com. Diakses: 20 November 2025. [Daring]. Tersedia pada: <https://www.gallup.com/analytics/349280/state-of-worlds-emotional-health.aspx>
- [9] K. Yang, T. Zhang, dan S. Ananiadou, "A mental state Knowledge-aware and Contrastive Network for early stress and depression detection on social media," *Inf. Process. Manag.*, vol. 59, no. 4, hlm. 102961, 2022, doi: 10.1016/j.ipm.2022.102961.
- [10] O. K. Burhan, "MELAMPAUI TRADISI : ANALISIS GALIAN TEKS UNTUK DATA KUALITATIF BERSKALA BESAR," vol. d, hlm. 162–168, 2024, doi: 10.24854/jpu1154.
- [11] A. Pathirana, D. K. Rajakaruna, D. Kasthurirathna, dan A. Atukorale, "Journal of Future Artificial Intelligence A Reinforcement Learning-Based Approach for Promoting Mental Health Using Multimodal Emotion Recognition," 2024.
- [12] J. B. Oluwagbemi, A. E. Mesioye, dan R. S. Akinbo, "Journal of Future Artificial Intelligence Depress-HybridNet : A Linguistic-Behavioral Hybrid Framework for Early and Accurate Depression Detection on Social Media," 2025.
- [13] E. Turcan, S. Muresan, dan K. McKeown, "Emotion-Infused Models for Explainable Psychological Stress Detection," *NAACL-HLT 2021 - 2021 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. Proc. Conf.*, hlm. 2895–2909, 2021, doi: 10.18653/v1/2021.naacl-main.230.
- [14] Y. Tolla, "Deteksi Stres dan Depresi Unggahan Media Sosial dengan Machine Learning," vol. 15, no. 1, hlm. 84–92, 2025.
- [15] M. E. Johan dan S. A. Azka, "G-Tech : Jurnal Teknologi Terapan using Multilayer Perceptron," vol. 7, no. 3, hlm. 958–966, 2023.
- [16] D. Suhartono, "ScienceDirect ScienceDirect Top 10 Countries with the Highest Rates of Stress," *Procedia Comput. Sci.*, vol. 216, hlm. 672–681, 2023, doi: 10.1016/j.procs.2022.12.183.
- [17] S. Muñoz dan C. A. Iglesias, "A text classification approach to detect psychological stress combining a lexicon-based feature framework with distributional representations," *Inf. Process. Manag.*, vol. 59, no. 5, hlm. 103011, 2022, doi: 10.1016/j.ipm.2022.103011.
- [18] A. Le Glaz *dkk.*, "Machine learning and natural language processing in mental health: Systematic review," *J. Med. Internet Res.*, vol. 23, no. 5, 2021, doi: 10.2196/15708.

- [19] R. KS, "Analyzing Online Conversations on Reddit: A Study of Stress and Anxiety Through Topic Modeling and Sentiment Analysis," *Cureus*, vol. 16, no. 9, 2024, doi: 10.7759/cureus.69030.
- [20] S. Ji, T. Zhang, L. Ansari, J. Fu, P. Tiwari, dan E. Cambria, "MentalBERT: Publicly Available Pretrained Language Models for Mental Healthcare," *2022 Lang. Resour. Eval. Conf. LREC 2022*, no. June, hlm. 7184–7190, 2022.
- [21] T. Zhang, "Natural language processing applied to mental illness detection : a narrative review," no. December, hlm. 1–13, 2021, doi: 10.1038/s41746-022-00589-7.
- [22] M. Islam, S. Hassan, S. Akter, dan F. Anam, "Healthcare Analytics A comprehensive review of predictive analytics models for mental illness using machine learning algorithms," *Healthc. Anal.*, vol. 6, no. June, hlm. 100350, 2024, doi: 10.1016/j.health.2024.100350.
- [23] L. Ilias, S. Mouzakitis, dan D. Askounis, "Calibration of Transformer-Based Models for Identifying Stress and Depression in Social Media," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 2, hlm. 1979–1990, 2024, doi: 10.1109/TCSS.2023.3283009.
- [24] G. F. Situmorang dan R. Purba, "Deteksi Potensi Depresi dari Unggahan Media Sosial X Menggunakan Teknik NLP dan Model IndoBERT," vol. 6, no. 2, hlm. 649–661, 2024, doi: 10.47065/bits.v6i2.5496.
- [25] M. R. Febriansyah, Nicholas, R. Yunanda, dan D. Suhartono, "Stress detection system for social media users," *Procedia Comput. Sci.*, vol. 216, hlm. 672–681, 2023, doi: 10.1016/j.procs.2022.12.183.
- [26] N. Oryngozha, P. Shamo, dan A. Igali, "Detection and Analysis of Stress-Related Posts in Reddit's Academic Communities," *IEEE Access*, vol. 12, hlm. 14932–14948, 2024, doi: 10.1109/ACCESS.2024.3357662.
- [27] J. Nicholas, S. Onie, dan M. E. Larsen, "Ethics and Privacy in Social Media Research for Mental Health.," *Curr. Psychiatry Rep.*, vol. 22, no. 12, hlm. 84, Nov 2020, doi: 10.1007/s11920-020-01205-9.
- [28] Z. Chen, R. Yang, S. Fu, N. Zong, H. Liu, dan M. Huang, "Detecting Reddit Users with Depression Using a Hybrid Neural Network SBERT-CNN," *Proc. - 2023 IEEE 11th Int. Conf. Healthc. Inform. ICHI 2023*, hlm. 193–199, 2023, doi: 10.1109/ICHI57859.2023.00035.
- [29] A. A. Dalimunthe dan N. A. Hasibuan, "Klasifikasi Arthropoda Dengan Pendekatan Convolutional Neural Network ( CNN )," vol. 5, no. 1, hlm. 119–125, 2025.
- [30] N. Yulianti, M. Afdal, M. Jazman, M. Megawati, dan A. Anofrizen, "Analisis Sentimen Masyarakat Menggunakan Algoritma Long Short Term Memory ( LSTM ) Pada Ulasan Aplikasi Halodoc," vol. 7, no. 2, 2025, doi: 10.47065/bits.v7i2.7243.