JITET (Jurnal Informatika dan Teknik Elektro Terapan)
Vol. 14 No. 3, pISSN: 2303-0577 eISSN: 2830-7062 http://dx.doi.org/10.23960/jitet.v14i3.9902

DETEKSI POTENSI DEPRESI PADA CUITAN X BERBAHASA INDONESIA MENGGUNAKAN INDOBERT DAN SUPPORT VECTOR MACHINE

Benedict Raditya Pradipta Ginting[1*], Yuyun Umaidah[2], Adhi Rizal[3]

[1,2,3]Universitas Singaperbangsa Karawang;Jl. HS.RonggoWaluyo, Puseurjaya, Telukjambe Timur Karawang, Jawa barat, Indonesia, 41361 ;Telp. (0267) 641177

Keywords:
Depression detection;
Feature extraction;
IndoBERT;
Social media X;
Support vector machine.

Corespondent Email:
ginting.youth@gmail.com

Abstrak. Kesehatan mental, khususnya depresi, merupakan isu kesehatan global yang seringkali sulit dideteksi secara dini. Media sosial X menjadi sarana potensial untuk deteksi dini berdasarkan ekspresi emosional pengguna. Penelitian ini bertujuan mengklasifikasikan cuitan berbahasa Indonesia ke dalam kategori depresi dan non-depresi menggunakan pendekatan gabungan IndoBERT sebagai pengekstraksi fitur dan Support Vector Machine (SVM) sebagai pengklasifikasi. Dataset terdiri dari 1.293 cuitan yang telah divalidasi oleh pakar psikologi klinis. Tahapan analisis meliputi pra-pemrosesan teks, penanganan ketidakseimbangan kelas menggunakan Back-translation dan SMOTE, ekstraksi fitur berdimensi padat dengan IndoBERT, dan klasifikasi menggunakan SVM berkernel linear. Hasil pengujian menunjukkan bahwa arsitektur gabungan ini sangat efektif, dengan tingkat akurasi mencapai 79,34% dan nilai recall untuk kelas depresi menyentuh 80% pada parameter optimal (C=0,01). Tingginya sensitivitas deteksi ini membuktikan bahwa integrasi representasi kontekstual IndoBERT dan ketangguhan margin klasifikasi SVM mampu mendeteksi indikasi gangguan kesehatan mental di ruang siber secara efisien. Hal ini berimplikasi pada kelayakan arsitektur tersebut untuk diimplementasikan sebagai landasan sistem peringatan dini (early warning system).

Abstract. *Mental health, particularly depression, is a global health issue that is often difficult to detect early. Social media X serves as a potential medium for early detection based on users' emotional expressions. This study aims to classify Indonesian tweets into depression and non-depression categories using a hybrid approach of IndoBERT as a feature extractor and Support Vector Machine (SVM) as a classifier. The dataset consists of 1,293 tweets validated by a clinical psychology expert. The analysis stages include text pre-processing, data imbalance handling using Back-translation and SMOTE, dense feature extraction with IndoBERT, and classification using a linear-kernel SVM. The experimental results show that this hybrid architecture is highly effective, achieving an accuracy of 79.34% and a recall value for the depression class reaching 80% at the optimal parameter (C=0.01). This high detection sensitivity proves that the integration of IndoBERT's contextual representation and SVM's classification margin robustness can efficiently detect indications of mental health disorders in cyberspace. This implies the feasibility of the architecture to be implemented as the foundation of an early warning system.*

1. PENDAHULUAN

Depresi adalah gangguan kesehatan mental yang berdampak signifikan terhadap kesejahteraan seseorang; sebuah laporan tahun 2025 dari *World Health Organization* mencatat bahwa lebih dari 332 juta orang di seluruh dunia menderita gangguan ini[1]. Di Indonesia, prevalensi depresi mencapai 3,7%, dengan sekitar 7,3 juta kasus, dan satu dari tiga remaja dilaporkan mengalami masalah kesehatan mental, termasuk kecemasan dan depresi [2]. Seiring dengan meningkatnya penggunaan media sosial, platform X, yang memiliki lebih dari 372 juta pengguna aktif bulanan, telah menjadi ruang terbuka bagi masyarakat untuk secara spontan mengekspresikan kondisi psikologis mereka. Sifat platform yang terbuka dan format teks tweet yang singkat menjadikan data dari X sebagai "cerminan psikologis digital" yang sangat relevan untuk mendeteksi kecenderungan depresi melalui analisis pola linguistik dan ekspresi emosional pengguna secara *real-time* [3], [4].

Meskipun media sosial menyediakan data teks yang melimpah, tantangan utama dalam mengklasifikasikan teks medis dan psikologis seperti deteksi depresi adalah ketidakseimbangan kelas yang parah. Jumlah cuitan organik yang benar-benar menunjukkan gejala depresi jauh lebih sedikit dibandingkan dengan cuitan biasa atau kelas mayoritas. Jika tidak ditangani, algoritma pembelajaran mesin akan menjadi bias dan gagal mengenali pola pada kelas minoritas. Oleh karena itu, teknik augmentasi data teks dan penyeimbangan fitur matematis diperlukan agar model dapat mendeteksi gejala depresi dengan sensitivitas yang tinggi.

Penelitian sebelumnya mengenai deteksi depresi di media sosial umumnya mengandalkan arsitektur *deep learning end-to-end*, seperti algoritma Bidirectional LSTM (BiLSTM) yang mencapai akurasi 94,12% [5], dan model LSTM-RNN yang mencapai metrik kinerja sebesar 86% [6]. Namun, model *Deep Learning* yang kompleks ini membutuhkan sumber daya komputasi yang sangat besar dan rentan terhadap kegagalan generalisasi atau overfitting ketika dihadapkan pada dataset berukuran terbatas. Di sisi lain, indikasi klinis depresi di media sosial jarang ditunjukkan oleh adanya kata kunci yang terisolasi, melainkan tercermin dalam struktur sintaksis, kontras emosional, dan konteks lengkap dari bahasa informal dalam satu postingan.

Untuk mengatasi tantangan dan celah (*gap*) penelitian tersebut, penelitian ini mengusulkan kebaruan metodologis dengan merancang pemisahan proses antara ekstraksi fitur (*feature extractor*) dan pengklasifikasian (*classifier*). Penelitian ini memanfaatkan IndoBERT, sebuah model bahasa mutakhir berbasis arsitektur *Bidirectional Encoder Representations from Transformers* (BERT) yang telah dilatih dengan korpus masif dan terbukti efektif dalam mengurai kompleksitas bahasa informal, kosakata gaul (slang), maupun struktur kalimat tidak baku di platform digital. IndoBERT difungsikan secara spesifik untuk mentransformasikan setiap teks cuitan menjadi sebuah representasi vektor padat (*dense vector embedding*) 768 dimensi yang mampu menangkap bobot semantik dan esensi kontekstual kalimat. Vektor numerik inilah yang kemudian diumpankan ke dalam algoritma *Support Vector Machine* (SVM) sebagai *classifier* akhir. Penerapan SVM dengan kernel linear sangat adaptif dan efisien untuk menemukan *hyperplane* pemisah paling optimal pada ruang fitur berdimensi padat, serta secara inheren memberikan ketahanan komputasi tinggi terhadap risiko *overfitting*. Efektivitas integrasi kedua pilar komputasi ini telah divalidasi oleh penelitian-penelitian sebelumnya dalam ranah analisis teks berbahasa Indonesia yang terbukti tangguh mencapai tingkat akurasi hingga 97% [7], [8].

Berdasarkan latar belakang dan analisis kesenjangan ini, penelitian ini bertujuan untuk menerapkan model hibrida yang mengintegrasikan ekstraksi fitur dari IndoBERT dengan klasifikasi SVM untuk mendeteksi potensi depresi dalam cuitan berbahasa Indonesia di platform X. Kinerja model hibrida ini akan dievaluasi secara menyeluruh menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* guna menghasilkan model komputasional yang tidak hanya akurat, tetapi juga memiliki landasan matematis yang kokoh dalam mengenali ekspresi emosional di balik dinamika bahasa digital.

2. TINJAUAN PUSTAKA

2.1 Ekstraksi Fitur Menggunakan IndoBERT

IndoBERT adalah model bahasa yang telah dilatih sebelumnya berdasarkan arsitektur Transformer, yang dikembangkan secara khusus untuk mengolah teks dalam bahasa Indonesia. Keunggulan utama IndoBERT terletak pada mekanisme *self-attention* dua arahnya, yang memungkinkannya menangkap konteks lengkap dan makna semantik dari kalimat. Dalam penelitian ini, IndoBERT tidak digunakan sebagai pengklasifikasi ujung ke ujung, melainkan secara khusus digunakan sebagai ekstraktor fitur. Model ini mengekstrak token [CLS] (Klasifikasi) dari lapisan tersembunyi terakhir untuk mengubah setiap cuitan menjadi sebuah representasi vektor padat berdimensi 768 [9].

2.2 Algoritma Support Vector Machine (SVM)

SVM adalah algoritma machine learning yang mengklasifikasikan data dengan mencari *hyperplane* optimal yang memaksimalkan margin antar kelas. SVM telah terbukti sangat tangguh dan efektif dalam memproses ruang fitur berdimensi tinggi yang dihasilkan oleh model bahasa besar seperti BERT. Penelitian ini menerapkan SVM dengan kernel linier, yang secara matematis sangat efisien dalam menemukan batas keputusan yang tepat dalam ruang fitur berdimensi 768 tanpa menimbulkan beban komputasi yang berlebihan. Model ini juga mengoptimalkan parameter regularisasi (C) dan menetapkan `*class_weight='balanced*' untuk memberlakukan penalti bobot yang lebih tinggi pada kesalahan klasifikasi untuk kelas minoritas.

2.3 Penanganan Ketidakseimbangan Data (Class Imbalance)

Untuk mengatasi ketidakseimbangan distribusi antara kelas cuitan depresi dan non-depresi, dua teknik augmentasi diterapkan:

1. *Back-translation*: Teknik augmentasi tingkat teks yang mensimulasikan terjemahan dua arah dengan menerjemahkan teks ke dalam bahasa perantara (Inggris) dan kemudian kembali ke bahasa Indonesia untuk menghasilkan variasi leksikal (parafrasa) tanpa mengubah makna emosional intinya [10].
2. SMOTE (*Synthetic Minority Over-sampling Technique*): Teknik augmentasi tingkat numerik yang bekerja dengan menghasilkan sampel fitur sintetis pada kelas minoritas berdasarkan kedekatan (*K-Nearest Neighbor*) dalam ruang vektor sehingga jumlahnya setara dengan kelas mayoritas [11].

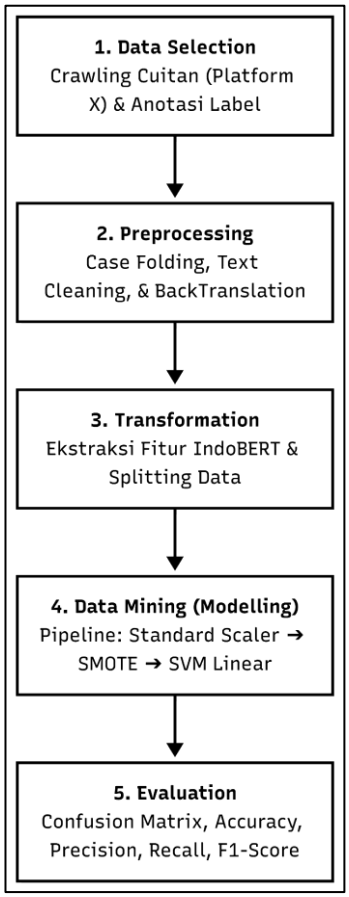
2.4 Penelitian Terkait

Pengembangan kerangka penelitian dalam studi ini didasarkan pada temuan berbagai penelitian sebelumnya. Penelitian yang dilakukan oleh Setyo Nugroho dkk. [5] dan Ivan Dwi Nugraha dkk. [6] menunjukkan bahwa *Deep Learning* murni (seperti BiLSTM dan LSTM-RNN) mendeteksi depresi dengan akurasi di atas 86%, namun model-model ini memerlukan komputasi yang sangat besar dan rentan terhadap *overfitting* pada dataset yang terbatas. Sebagai solusi yang lebih stabil, penelitian oleh Wijaya dkk. [12] menunjukkan bahwa penggabungan BERT dan SVM dapat meningkatkan efisiensi dan akurasi klasifikasi hingga 93,49%. Khusus untuk analisis teks bahasa Indonesia, studi komparatif oleh Hidayatulloh dkk. [7] dan Irianti dkk. [8] memvalidasi bahwa integrasi IndoBERT (sebagai ekstraktor fitur) dan SVM (sebagai pengklasifikasi) sangat efektif dalam memetakan batas klasifikasi teks sentimen yang kompleks, dengan mencapai akurasi keseluruhan sebesar 97%.

3. METODE PENELITIAN

3.1. Rancangan Penelitian

Penelitian ini merupakan proyek penelitian kuantitatif yang didasarkan pada klasifikasi teks prediktif menggunakan kerangka kerja Knowledge Discovery in Databases (KDD). Arsitektur pemodelan dirancang menggunakan pendekatan hibrida, di mana proses pemahaman konteks linguistik ditangani oleh model Transformer, sedangkan klasifikasi kategori sentimen ditangani secara matematis oleh algoritma pembelajaran mesin tradisional. Alur kerja penelitian secara keseluruhan ditunjukkan pada Gambar 1.



Gambar 1 Rancangan Penelitian

3.2. Sumber dan Partisipan Data

Sumber data utama yang digunakan adalah teks cuitan publik dalam bahasa Indonesia dari platform media sosial X. Kumpulan data yang diekstraksi berfokus pada postingan dari tanggal 1 Januari 2025 hingga 31 Desember 2025. Untuk pelabelan *ground-truth*, penelitian ini menghindari metode *crowdsourcing* umum guna mencegah gangguan label pada topik kesehatan mental yang bersifat subjektif. Anotasi data dilakukan secara eksklusif oleh Kayla Nurisa untuk memastikan konsistensi dan akurasi dalam pelabelan kelas "Depresi" dan "Non-Depresi".

3.3. Teknik Pengumpulan Data

Pengumpulan data dilakukan dengan menggunakan teknik *web scraping* melalui alat *Tweet-Harvest* v2.6.1, yang dijalankan pada ekosistem Python di Google Colaboratory. Kueri pencarian dibatasi pada parameter `lang:id` dengan menggunakan kombinasi kata kunci emosional tertentu: 'depresi', 'kecewa', dan 'menyerah'. Proses ini menghasilkan 3.105 baris data mentah, yang setelah melalui proses pembersihan duplikasi dan nilai nol, berkurang menjadi 2.792 baris. Setelah diberi anotasi dan disaring oleh para ahli untuk menghilangkan konteks yang ambigu, diperoleh kumpulan data akhir sebanyak 1.293 cuitan (918 Non-Depresi dan 375 Depresi).

3.4. Teknik Analisis Data

Teknik analisis data dilakukan secara berurutan, dengan fokus pada pengoptimalan prapemrosesan, ekstraksi fitur, dan klasifikasi:

3.4.1. Prapemrosesan Teks (Preprocessing)

Teks diubah menjadi huruf kecil (*case folding)* dan dibersihkan dari gangguan (URL, @mentions, tagar, dan emoji). Teknik *stemming* dan *Stopword Removal* secara eksplisit tidak digunakan untuk mempertahankan struktur sintaksis yang dibutuhkan oleh model bahasa. Ketidakseimbangan kelas diatasi pada tingkat teks menggunakan teknik *back-translation* berdasarkan model *Helsinki-NLP/opus-mt* (id-en dan en-id) untuk merumuskan ulang 375 titik data kelas Depresi menjadi 750 titik data sintetis yang valid.

3.4.2. Transformasi Fitur (Transformation)

Teks yang telah dibersihkan dimasukkan ke dalam model *indobenchmark/indobert-base-p2* (tanpa pembedaan huruf besar-kecil). Fitur diekstraksi dari token [CLS] pada *last hidden state*, mengubah setiap tweet menjadi matriks vektor padat berdimensi 768. Data numerik ini kemudian dibagi (80% data pelatihan, 20% data uji) [13].

3.4.3. Pemodelan (Data Mining)

Analisis klasifikasi dilakukan melalui *pipeline* terintegrasi pada data latih menggunakan *library imblearn.pipeline*. Proses dalam *pipeline* meliputi:

1. *Standard Scaler*: Menstandarkan distribusi vektor fitur ke rata-rata 0 dan standar deviasi 1.
2. *SMOTE*: Penyeimbangan dimensi fitur sintetis melalui *K-Nearest Neighbor* pada kelas minoritas dalam data latih.
3. *SVM*: Implementasi klasifikasi menggunakan algoritma *Support Vector Machine* dengan parameter *kernel='linear', class_weight='balanced'*, dan *probability=True.*

3.4.4. Optimasi dan Evaluasi

Penyesuaian *hyperparameter* dilakukan menggunakan *GridSearchCV* (rentang C = [0,01, 0,1, 1, 5, 10, 50, 100]) dengan skema *5-Fold Cross-Validation* yang dioptimalkan untuk *Recall Macro*. Kinerja model akhir diukur menggunakan matriks kebingungan (*Confusion Matrix*) pada 20% data uji berdasarkan metrik Akurasi, Presisi, *Recall*, dan *F1-Score* [14].

4. HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

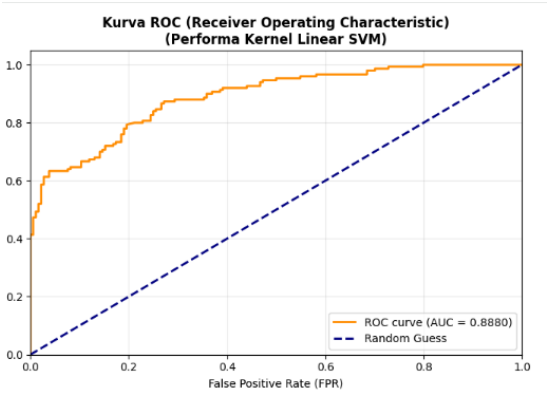
Berdasarkan pemodelan yang dilakukan, matriks vektor berdimensi 768 yang diekstraksi dari IndoBERT diproses menggunakan algoritma *Support Vector Machine* (SVM). Dari hasil pencarian parameter menggunakan skema *5-Fold Cross-Validation*, konfigurasi paling optimal untuk mengklasifikasikan data adalah dengan menggunakan nilai regularisasi C=0,01. Pemodelan ini juga melibatkan penerapan Standard Scaler untuk menormalkan fitur dan SMOTE untuk menyeimbangkan kelas dalam data latih.

Model terbaik ini kemudian dievaluasi menggunakan 334 titik data uji (20% dari total data) yang telah disisihkan sebelumnya. Evaluasi kuantitatif menunjukkan bahwa model mencapai akurasi 79,34%, presisi 79,16%, *recall* 79,40%, dan *F1-score* sebesar 79,22%.

Tabel 1 Hasil Evaluasi Performa Model IndoBERT-SVM

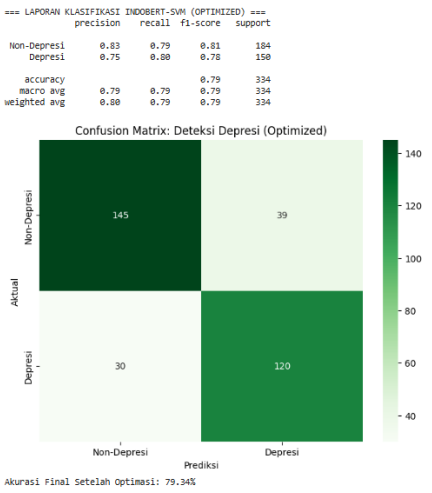
Metrik Evaluasi	Skor
Accuracy	0.7934
Precision	0.7916
Recall	0.7940
F1-Score	0.7922

Kemampuan model dalam membedakan antar kelas juga diukur menggunakan kurva *Receiver Operating Characteristic* (ROC), seperti yang ditunjukkan pada Gambar 2.



Gambar 2 Kurva ROC Model IndoBERT-SVM

Nilai *Area Under Curve* (AUC) sebesar 0,8880 menunjukkan bahwa model tersebut memiliki probabilitas sebesar 88,8% untuk secara akurat mengidentifikasi cuitan yang mengindikasikan depresi. Untuk menganalisis distribusi prediksi secara lebih rinci, *Confusion Matrix* pada Gambar 3 memperlihatkan perbandingan antara nilai aktual dan hasil prediksi model.



Gambar 3 *Confusion Matrix* Klasifikasi pada Data Uji

Berdasarkan matriks tersebut, model tersebut berhasil mengklasifikasikan 145 titik data ke dalam kelas Non-Depresi dan 120 titik data ke dalam kelas Depresi dengan benar. Kesalahan klasifikasi (klasifikasi yang salah) terjadi pada 39 kasus positif palsu dan 30 kasus negatif palsu.

4.2 Pembahasan

Secara keseluruhan, integrasi IndoBERT dan SVM terbukti efektif dalam mendeteksi potensi depresi, terutama dengan tingkat *recall* untuk kelas depresi yang mencapai 80% (0,8000) dalam kondisi optimal. Tingkat *recall*

yang tinggi ini sejalan dengan tujuan utama deteksi dini kesehatan mental, di mana sistem dirancang untuk meminimalkan jumlah *false negative* (kegagalan mendeteksi individu yang terkena dampak) sehingga mereka yang membutuhkan bantuan tidak terlewatkan.

Peningkatan sensitivitas model ini merupakan hasil langsung dari penerapan teknik-teknik untuk mengatasi ketidakseimbangan data. Perbandingan kinerja masing-masing skenario ditunjukkan pada Tabel 2.

Tabel 2 Perbandingan Skenario Penanganan Data terhadap Performa Model

Skenario	Accuracy	Recall Kelas Depresi	Keterangan
Baseline	0.7220	0.3467	Bias terhadap kelas Mayoritas
Baseline + SMOTE	0.7297	0.6667	Model mulai memperhatikan kelas minoritas namun belum optimal
Baseline+ SMOTE+ StandardScaler	0.7375	0.6933	Kenaikan pada *recall,* tetapi masih belum cukup bisa menangkap kompleksitas bahasa
Backtranslation	0.8114	0.7733	Akurasi tertinggi, model lebih sensitive terhadap data.
Backtranslation + SMOTE	0.7695	0.7533	Terjadi penurunan performa pada akurasi dan recall.
Backtranslation + SMOTE + StandardScaler	0.7934	0.8000	Nilai recall paling tinggi, dan akurasi tidak timpang.

Sebagaimana ditunjukkan pada Tabel 2, model *baseline* (tanpa penyeimbangan data) cenderung bias ke arah kelas mayoritas, dengan nilai *recall* hanya sebesar 0,3467. Penerapan *Back-translation*, SMOTE, dan *Standard Scaler* secara bersamaan terbukti membantu model membentuk batas keputusan yang lebih seimbang. Pemilihan parameter C=0,01 juga menunjukkan bahwa SVM dengan *soft margin* cukup adaptif dalam menangani variasi struktur kalimat informal di platform media sosial X.

Dibandingkan dengan penelitian sebelumnya yang menggunakan model *deep learning end-to-end* seperti BiLSTM oleh Setyo Nugroho dkk [5], pendekatan *hybrid* ini menawarkan alternatif yang jauh lebih efisien secara komputasi dan tidak rentan terhadap *overfitting*. Hasil penelitian ini juga mengonfirmasi temuan Oliveira dkk. [15] dan Hidayatulloh dkk. [7], yang menunjukkan bahwa SVM mampu menghasilkan klasifikasi yang stabil saat memanfaatkan pengkodean padat dari ekstraktor fitur berbasis Transformer.

Meskipun performa model cukup memadai, analisis kesalahan (*error analysis*) menunjukkan adanya keterbatasan algoritma dalam memahami konteks kalimat yang ambigu. Contoh analisis kesalahan dapat dilihat pada Tabel 3.

Tabel 3 Contoh Analisis Kesalahan Prediksi (False Positive & False Negative)

No.	Cuitan	Kategori Kesalahan	Keterangan
1.	aku tau kamu capek tapi aku tau kamu wanita hebat aku yakin kamu ga akan menyerah dengan laki seperti ini tapi laki seperti ini pasti meninggalkan luka dan trauma yang cukup dalam bagimu kan	Ambiguitas Subjek Pengalaman (Konteks Empati)	Teks ini merupakan bentuk dukungan moral (empati) kepada pihak kedua ("kamu"). Namun, model IndoBERT menangkap akumulasi kata berbobot emosi negatif yang sangat pekat (*"capek", "luka", "trauma", "menyerah"*). Model gagal membedakan antara subjek yang menderita dengan subjek yang memberikan simpati, sehingga *hyperplane* SVM terdorong ke arah kelas depresi.
2.	sender lagi skripsian lop u skripsi terus bingung pingin cepet lulus tapi juga takut lulus mending gimana ya kadang kepikiran apa enak kali ya kalau genre hidup jadi survival tapi pasti menyerah duluan	Hiperbola Keluh Kesah Akademik (*Burnout*)	Teks merepresentasikan stres akademik atau *burnout* mahasiswa. Penggunaan hiperbola *"genre hidup jadi survival"* dan leksikon keputusasaan *"menyerah duluan"* diinterpretasikan oleh model sebagai hilangnya harapan hidup (gejala depresi). Model belum sepenuhnya mampu memisahkan antara stres situasional akademik dengan gangguan mental klinis.
3.	ku ngerasa lowkey depresi di	Penggunaan Leksikon	Kemunculan kata kunci eksplisit *"depresi"* (dalam

	semester sih bukan ga ngumpul tugas tp gatau gimana mau lanjutin skripsi amp stuck krn udh g cocok sama judulnya akhirnya continue dg ngambil decision ganti judul skripsi dan start dari awal even if diomel omelin dosen u just gotta face ur fear	Klinis Kasual	frasa *slang "lowkey depresi"*) memberikan pembobotan fitur yang sangat tinggi pada model. Meskipun paruh akhir kalimat menunjukkan sikap resilien atau bangkit dari masalah (*"face ur fear"*), SVM telah terlanjur memprioritaskan kata klinis tersebut di awal kalimat sebagai sinyal utama kelas minoritas.
4.	media media ternama hanya menyiarkan berita berita kerusakan akibat demonstrasi tidak lebih jujur saya kecewa kecewa kecewa demonstrasi seolah menjadi gerakan negatif yang bersifat anarkis dan vandalis seperti hanya membuat kericuhan di masyarakat padahal tidak	Sentimen Negatif Eksternal (Kemarahan Publik)	Cuitan ini sarat akan sentimen negatif ekstrem (pengulangan kata *"kecewa"*, *"kerusakan"*, *"anarkis"*). Kesalahan klasifikasi terjadi karena model menyamakan *polaritas sentimen negatif* (kemarahan terhadap pihak eksternal/media) dengan *kondisi depresi* (perasaan negatif yang diarahkan ke dalam diri sendiri).
5.	nih fandom depresi ngemis ngemis pengen masuk koz tapi gak suka karena ada si hyberuk heii kayak koz mau aja sama idol lu even kalo koz bukan bagian beruk juga kagak mungkin idol lo dipungut	Konteks Spesifik Fanwar & *Harassment*	Teks ini adalah pertikaian antarpenggemar (*fanwar*) yang sangat spesifik pada kultur X. Kata *"depresi"* di sini digunakan sebagai kata sifat untuk merendahkan kelompok tertentu (*"fandom depresi"*), ini adalah bukan merujuk pada kondisi mental seseorang. Model gagal memahami konteks semantik dari *slang* kultural ini dan terjebak pada deteksi leksikal mentah.

Dalam kasus hasil *False Positive*, kesalahan prediksi umumnya dipicu oleh penggunaan bahasa yang berlebihan terkait kelelahan sehari-hari (seperti *burnout* akibat tugas akademis) atau penggunaan kata kunci negatif yang tidak merujuk pada kondisi psikologis. Sementara itu, dalam kasus hasil *False Negative*, model masih kesulitan mendeteksi gejala depresi yang diungkapkan secara implisit atau disamarkan oleh humor gelap (*dark humor*).

Temuan ini menunjukkan bahwa penelitian lebih lanjut masih diperlukan, yang mencakup pengembangan modul untuk pemahaman pragmatik atau deteksi sarkasme. Secara praktis, kemampuan model ini sudah memadai untuk digunakan sebagai prototipe sistem peringatan dini di platform digital.

5. KESIMPULAN

Berdasarkan hasil pengujian dan analisis model gabungan IndoBERT dan Support Vector Machine (SVM) untuk mendeteksi potensi depresi dalam cuitan di media sosial X, dapat ditarik kesimpulan sebagai berikut:

a. Hasil yang diperoleh: Penerapan arsitektur *hybrid* yang menggunakan IndoBERT sebagai ekstraktor fitur dan SVM kernel linier sebagai klasifikasi terbukti sangat efektif. Dengan konfigurasi optimal (parameter C=0,01, penerapan augmentasi back-translation, penyeimbangan SMOTE, dan *Standard Scaler*), model ini mencapai akurasi sebesar 79,34% dan sensitivitas (*recall*) sebesar 80% (0,8000) untuk kelas depresi.
b. Kelebihan: Pendekatan *hybrid* ini menawarkan efisiensi komputasi yang tinggi dan stabilitas yang lebih baik pada kumpulan data berskala kecil hingga menengah dibandingkan dengan model *Deep Learning end-to-end*, yang rentan terhadap *overfitting*. Tingkat *recall* yang tinggi menunjukkan bahwa model ini memiliki sensitivitas klinis yang sangat baik dalam meminimalkan tingkat *false negative*, sehingga sangat cocok untuk diterapkan sebagai landasan sistem peringatan dini.
c. Kekurangan: Keterbatasan utama model ini terletak pada ketidakmampuannya untuk mengolah ambiguitas linguistik tingkat tinggi atau penalaran semantik yang kompleks. Algoritma ini masih rentan terhadap kesalahan prediksi ketika dihadapkan pada cuitan yang menggunakan humor gelap (*dark humor*), sarkasme

ekstrem, hiperbola situasional (seperti tekanan akademik), dan teks yang mengungkapkan gejala secara sangat implisit.

d. Kemungkinan Pengembangan Selanjutnya: Untuk mengatasi keterbatasan-keterbatasan ini, penelitian selanjutnya disarankan untuk mengintegrasikan modul pendeteksi sarkasme pra-klasifikasi atau menerapkan ekstraksi fitur berbasis aturan dengan menggunakan leksikon psikologi klinis. Penggunaan *Large Language Models* (LLM) generatif dapat dieksplorasi lebih lanjut untuk meningkatkan kemampuan sistem dalam memahami pragmatik dan konteks emosional di lingkungan media sosial yang sangat dinamis. Selain itu, penelitian selanjutnya disarankan untuk melibatkan lebih dari satu ahli psikologi sebagai anotator dalam proses pelabelan data, serta mengukur tingkat kesepakatan di antara para anotator menggunakan metrik Inter-Annotator Agreement (IAA) seperti *Cohen's Kappa* atau *Fleiss' Kappa*. Pendekatan ini akan meningkatkan objektivitas dan kredibilitas label data yang dihasilkan, sehingga menjadikan dataset tersebut sebagai referensi yang lebih kuat untuk penelitian selanjutnya mengenai deteksi depresi dalam bahasa Indonesia.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Singaperbangsa Karawang, secara khusus kepada Koordinator Program Studi Informatika beserta dosen pembimbing yang telah memberikan arahan, dukungan, dan fasilitas selama proses penyusunan penelitian ini. Penghargaan juga diberikan kepada pakar psikologi yang telah membantu proses anotasi dataset secara objektif.

DAFTAR PUSTAKA

[1] World Health Organization, "Depressive disorder (depression)," 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[2] Center for Reproductive Health, University of Queensland, and Johns Hopkins Bloomberg School of Public Health, "Indonesia - National Adolescent Mental Health Survey (I-NAMHS)," Oct. 2022.

[3] K. K. Putri and E. B. Setiawan, "Depression Detection in Indonesian X Social Media Text using Convolutional Neural Networks and Long Short-Term Memory with TF- IDF and FastText Methods," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 2, pp. 557–574, Apr. 2025, doi: 10.52436/1.jutif.2025.6.2.4206.

[4] Y. Tolla and Kusrini, "Deteksi Stres dan Depresi Unggahan Media Sosial dengan Machine Learning," *JURNAL FASILKOM*, vol. 15, no. 1, pp. 84–92, Apr. 2025.

[5] K. S. Nugroho, I. Akbar, A. N. Suksmawati, and Istiada, "DETEKSI DEPRESI DAN KECEMASAN PENGGUNA TWITTER MENGGUNAKAN BIDIRECTIONAL LSTM," *The 4th Conference on Innovation and Application of Science and Technology (CIASTECH 2021)*, Dec. 2021.

[6] I. D. Nugraha and Y. Azhar, "Deteksi Depresi Pengguna Twitter Indonesia Menggunakan LSTM-RNN," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 3, pp. 320–329, Dec. 2022, doi: 10.23887/janapati.v11i3.50674.

[7] S. Hidayatulloh, L. Muflikhah, and R. S. Perdana, "Implementasi Embedding IndoBERT dan Support Vector Machine (SVM) Untuk Analisis Sentimen Publik terhadap Layanan Biznet," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 9, pp. 2548–964, Sep. 2025, [Online]. Available: http://j-ptiik.ub.ac.id

[8] A. Irianti, Halimah, Sutedi, and M. Agariana, "Integration of BERT and SVM in Sentiment Analysis of Twitter/X Regarding Constitutional Court Decision No. 60/PUU-XXII/2024," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 2, pp. 469–482, Apr. 2025, doi: 10.52436/1.jutif.2025.6.2.4068.

[9] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings ofNAACL-HLT 2019*, vol. 1, pp. 4171–4186, Jun. 2019, doi: https://doi.org/10.18653/v1/N19-1423.

[10] I. A. Rahma and L. H. Suadaa, "Penerapan Text Augmentation untuk Mengatasi Data yang Tidak Seimbang pada Klasifikasi Teks Berbahasa Indonesia," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 6, pp. 1329–1340, Dec. 2023, doi: 10.25126/jtiik.2023107325.

[11] M. Sulistiyono, Y. Pristyanto, S. Adi, and G. Gumelar, "Implementasi Algoritma Synthetic Minority Over-Sampling Technique untuk Menangani Ketidakseimbangan Kelas pada Dataset Klasifikasi," *SISTEMASI: Jurnal Sistem Informasi*, vol. 10, no. 2, pp. 445–459,

May 2021, doi: https://doi.org/10.32520/stmsi.v10i2.1303.
[12] H. Wijaya, M. F. Hadi, and N. Sulistianingsih, "Using Sentiment Analysis with BERT and SVM for Detect Mental Health Detection on Social Media," *Journal of Digital Health Innovation and Medical Technology*, vol. 1, no. 2, Feb. 2025.
[13] V. R. Joseph, "Optimal ratio for data splitting," *Stat. Anal. Data Min.*, vol. 15, no. 4, pp. 531–538, Aug. 2022, doi: 10.1002/sam.11583.
[14] K. R. Amaliah, U. Enri, and So. Defiyanti, "ANALISIS SENTIMEN PENGGUNA WEBSITE SISKA MENGGUNAKAN ALGORITMA NAIVE BAYES DENGAN OPTIMASI PARTICLE SWARM OPTIMIZATION," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.6839.
[15] L. L. Oliveira, X. Jiang, A. N. Babu, P. Karajagi, and A. Daneshkhah, "Effective Natural Language Processing Algorithms for Gout Flare Early Alert from Chief Complaints," Nov. 2023, doi: 10.1101/2023.11.28.23299150.