E-ISSN 2963-9174

Indonesian Journal of Computer Science Research
Journal page: https://subset.id/index.php/IJCSR

Deteksi Indikasi Gangguan Kesehatan Mental Berbasis Teks Menggunakan NLP Dengan Teknik Augmentasi EDA

Mental Health Disorder Indications Detection Based on Text Using NLP with EDA Augmentation Techniques

Sherly Dian Tiara [1,a)], Erna Daniati [2,b)], Arie Nugroho [3,c)]

[1, 2, 3)] Program Studi Sistem Informasi , Fakultas Teknik dan Ilmu Komputer,
Universitas Nusantara PGRI Kediri, Jl. KH. Achmad Dahlan No.76, Kediri 64112, Indonesia

Author Emails
[a)] *Corresponding author: Erna Daniat, Email: ernadaniati@unpkediri.ac.id*

[1)]Jurusan Matematika dan Ilmu Pengetahuan Alam, 2) Jurusan Elins
Universitas Gadjah Mada, Jl. Bulaksumur, Yogyakarta 55281, Indonesia

Received: 2 Juli 2026; Accepted: 27 Juli 2026

KEYWORDS

Kata kunci:
Deteksi Dini
Easy Data Augmentation
Logistic Regression
Natural Language Processing
Kesehatan Mental

Keywords:
Early Detection
Easy Data Augmentation
Logistic Regression
Natural Language Processing
Mental Health

ABSTRAK

Abstrak. Penelitian ini bertujuan membangun model klasifikasi penyaringan awal (*screening*) gangguan kesehatan mental dari data teks media sosial menggunakan kerangka kerja CRISP-DM. Masalah utama berupa ketidakseimbangan jumlah data antar kategori diatasi menggunakan teknik *Easy Data Augmentation* (EDA). Algoritma *Logistic Regression* dan ekstraksi fitur TF-IDF digunakan untuk mengklasifikasikan enam kategori kondisi mental. Hasil pengujian menunjukkan bahwa model dengan EDA mengalami sedikit penurunan akurasi global menjadi 0.74 (dibandingkan 0.76 tanpa EDA), namun berhasil meningkatkan *Recall* pada kelas minoritas *Mental illness* secara signifikan dari 0.28 menjadi 0.56. Peningkatan ini membuktikan bahwa EDA efektif memperkaya variasi linguistik pada data yang terbatas. Model telah divalidasi oleh pakar psikolog dan diimplementasikan ke dalam aplikasi berbasis web sebagai alat bantu deteksi dini indikatif, bukan sebagai diagnosis medis klinis.

Abstract. *This study aims to build a classification model for the early screening of mental health disorders from social media text data using the CRISP-DM framework. The primary issue of data imbalance between categories was addressed using the Easy Data Augmentation (EDA) technique. Logistic Regression algorithm and TF-IDF feature extraction were used to classify six categories of mental conditions. Test results showed that the model with EDA experienced a slight decrease in global accuracy to 0.74 (compared to 0.76 without EDA) but successfully increased the Recall for the minority class, Mentalillness, significantly from 0.28 to 0.56. This improvement proves that EDA effectively enriches linguistic variation in limited data. The model has been validated by a psychologist and implemented into a web-based application as an indicative early detection tool, not a clinical medical diagnosis.*

1. Pendahuluan

Gangguan kesehatan mental merupakan isu kesehatan masyarakat yang signifikan karena berdampak langsung pada kualitas hidup, produktivitas, serta relasi sosial individu[1]. Seiring dengan perkembangan era digital, platform media sosial telah menjadi wadah bagi masyarakat untuk mengekspresikan emosi, keluh-kesah,

dan pengalaman hidup mereka secara terbuka[2]. Fenomena ini membuka peluang besar bagi pemanfaatan pendekatan komputasional, khususnya Natural Language Processing (NLP), untuk melakukan pemantauan dan deteksi dini terhadap sinyal gangguan mental melalui jejak digital berbasis teks[3]. Deteksi sedini mungkin sangat diperlukan agar bantuan dapat segera diberikan sebelum kondisi psikologis seseorang berkembang menjadi lebih berat[4].

Meskipun dataset internasional seperti dari platform Reddit [5] sering digunakan dalam riset karena skalanya yang besar, tantangan utama yang sering dihadapi adalah masalah ketidakseimbangan kelas (class imbalance). Distribusi label yang tidak merata, di mana kategori tertentu lebih dominan dibandingkan kategori lainnya, dapat menyebabkan model cenderung bias dan memiliki performa yang kurang optimal saat menghadapi variasi bahasa yang nyata[6]. Selain itu, kurangnya variasi dalam ekspresi bahasa pada kategori minoritas membuat model sulit mengenali pola linguistik yang beragam. Keterbatasan data ini menjadi hambatan serius dalam membangun sistem klasifikasi yang adil dan akurat untuk berbagai jenis kondisi mental[7].

Untuk menjawab tantangan tersebut, penelitian ini menerapkan teknik Easy Data Augmentation (EDA) guna memperkaya variasi data latih tanpa harus mengumpulkan data baru dari sumber luar[8]. Teknik EDA yang terdiri dari operasi penggantian sinonim, penyisipan acak, pertukaran acak, dan penghapusan acak terbukti efektif dalam meningkatkan kinerja klasifikasi pada dataset yang berukuran terbatas[9]. Fokus utama penelitian ini adalah menyiapkan dataset publik dari Reddit menjadi data yang terstruktur melalui proses pembersihan dan praproses teks yang ketat. Tujuan akhirnya adalah melakukan evaluasi mendalam terhadap dampak augmentasi EDA pada model baseline menggunakan metrik evaluasi standar seperti akurasi, precision, recall, dan F1-score.

Kontribusi penting dari makalah ini adalah penyediaan dokumentasi dan pipeline pengolahan data yang dapat menjadi pedoman dalam membangun sistem penyaringan (screening) awal untuk risiko kesehatan mental. Perlu ditegaskan bahwa penelitian ini memiliki batasan pada penggunaan data publik tanpa melibatkan uji klinis medis secara langsung[10]. Hasil prediksi yang dihasilkan oleh sistem ini bersifat indikatif sebagai bantuan teknologi dan mutlak tidak memiliki validitas sebagai diagnosis medis klinis. Seluruh proses penegakan diagnosis resmi tetap merupakan wewenang tenaga profesional medis seperti dokter atau psikolog terkait.

2. Tinjauan Pustaka

Deteksi dini gangguan mental melalui pendekatan komputasional merupakan upaya klasifikasi risiko berdasarkan jejak digital untuk mengenali indikasi awal dari kondisi psikologis tertentu. Dalam pelaksanaannya, Natural Language Processing (NLP) bertindak sebagai cabang kecerdasan buatan utama yang digunakan untuk memproses, memahami, dan mengekstraksi representasi fitur dari teks yang dipublikasikan oleh pengguna[11]. Pengolahan NLP pada penelitian ini difokuskan pada ekstraksi fitur dan pembangunan model dasar (baseline) untuk penyaringan kesehatan mental[12].

Representasi teks sering kali dilakukan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) yang memberikan bobot pada kata berdasarkan seberapa sering kata tersebut muncul dalam sebuah dokumen, namun jarang muncul pada dokumen lain[13]. Fitur TF-IDF ini kemudian diproses menggunakan algoritma Logistic Regression, yakni metode klasifikasi yang memodelkan probabilitas suatu kelas berdasarkan kombinasi linear fitur input[14]. Kombinasi TF-IDF dan Logistic Regression sangat efisien untuk data berukuran besar, mudah dijelaskan, dan terbukti kuat sebagai model baseline dalam klasifikasi teks[15]. Untuk mengatasi masalah data terbatas, digunakan teknik augmentasi teks Easy Data Augmentation (EDA) yang terdiri dari empat operasi: penggantian sinonim (synonym replacement), penyisipan acak (random insertion), pertukaran acak (random swap), dan penghapusan acak (random deletion)[16][17].

Terkait kajian penelitian terdahulu,[8] menunjukkan bahwa penerapan teknik EDA secara signifikan mampu meningkatkan performa klasifikasi teks, terutama pada dataset yang berukuran kecil karena model dapat mengenali lebih banyak variasi bentuk kalimat. Dalam konteks klasifikasi kesehatan mental, Inamdar dkk. [6] berhasil memperoleh skor F1 sebesar 0,76 dalam mendeteksi stres dari postingan platform Reddit menggunakan pendekatan Machine Learning konvensional seperti TF-IDF[18]. Sejalan dengan itu, Ayyalasomayajula [5] membuktikan bahwa penggunaan regresi logistik dengan fitur TF-IDF mampu memprediksi depresi secara komputasional. Penelitian lain dari Nugroho dkk. [19] dan Maldini dkk. [20] juga menyoroti pentingnya

penanganan data yang tidak seimbang (imbalance data) menggunakan metode seperti SMOTE dan klasifikasi stacking guna mendongkrak akurasi deteksi indikasi psikologis di media sosial.

3. Metode penelitian

Penelitian ini menggunakan kerangka CRISP-DM (Cross-Industry Standard Process for Data Mining) karena tahapannya jelas dan cocok untuk penelitian klasifikasi teks berbasis data [21]. Tahapan CRISP-DM terdiri dari: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment.

Gambar 1. Alur Penelitian

3.1. Business Understanding

Fase pertama, *Business Understanding*, diawali dengan melakukan identifikasi mendalam terhadap fenomena peningkatan prevalensi gangguan kesehatan mental yang berbanding lurus dengan kecenderungan individu dalam menggunakan platform digital sebagai proksi ekspresi emosional . Tujuan utama dari tahapan ini adalah merumuskan kerangka kerja komputasional untuk membangun sebuah instrumen penyaringan (*screening*) awal yang mampu mengenali indikasi gangguan mental melalui analisis pola bahasa pada teks. Fokus penelitian diarahkan pada evaluasi efektivitas teknik *Natural Language Processing* (NLP) dalam mengoptimalkan performa klasifikasi, khususnya dalam menghadapi tantangan keterbatasan variasi data. Penting untuk ditegaskan bahwa sasaran luaran dari fase ini adalah pengembangan alat pendukung keputusan yang bersifat indikatif, sehingga sistem ini mutlak tidak memiliki otoritas atau validitas sebagai diagnosis medis klinis

3.2 Data Understanding

Fase kedua, *Data Understanding*, melibatkan eksplorasi sistematis terhadap korpus data publik yang bersumber dari platform Reddit melalui repositori Kaggle. Proses ini mencakup audit terhadap struktur dataset yang terdiri dari beberapa atribut utama, yaitu *title* (judul), *selftext* (isi postingan), serta label kategori (*subreddit*) sebagai variabel target klasifikasi. Peneliti melakukan asesmen terhadap integritas data dengan mengidentifikasi berbagai anomali, seperti nilai kosong (*missing values*) serta entri teks non-informatif yang berlabel '[removed]' atau '[deleted]'. Selain itu, dilakukan analisis statistik terhadap distribusi label untuk memetakan derajat ketidakseimbangan kelas (*class imbalance*). Hasil observasi menunjukkan adanya dominasi kategori tertentu (seperti BPD) berbanding kategori minoritas (seperti *Mentalillness*), yang secara teoretis berpotensi menimbulkan bias pada proses pembelajaran model jika tidak ditangani dengan tepat . Pemahaman komprehensif terhadap karakteristik data ini menjadi fondasi krusial bagi tahap persiapan data (*Data Preparation*) selanjutnya.

3.3. Data Preparation

Fase ketiga, *Data Preparation*, merupakan tahapan paling krusial untuk menyiapkan data agar siap dilatih oleh mesin. Proses ini meliputi pembersihan teks (*cleaning*) dari karakter non-alfanumerik dan URL, normalisasi huruf menjadi huruf kecil (*lowercase*), serta penggabungan atribut *title* dan *selftext* menjadi satu fitur teks tunggal yang lebih kaya informasi. Setelah dataset dibagi menjadi data latih dan data uji secara proporsional menggunakan metode *stratified split*, teknik *Easy Data Augmentation* (EDA) diterapkan secara eksklusif pada data latih di kelas-kelas minoritas. Empat operasi EDA yang diaplikasikan meliputi penggantian sinonim (*synonym replacement*),

penyisipan acak (*random insertion*), pertukaran acak (*random swap*), dan penghapusan acak (*random deletion*) guna memperkaya keragaman ekspresi linguistik.

3.4. Modeling

Fase keempat, *Modeling*, dilakukan dengan mentransformasikan data teks yang telah dibersihkan dan diaugmentasi menjadi representasi vektor numerik menggunakan metode ekstraksi fitur *Term Frequency-Inverse Document Frequency* (TF-IDF). Algoritma *Logistic Regression* kemudian diterapkan untuk membangun dua skenario model klasifikasi sebagai bahan perbandingan: model *baseline* pertama dilatih menggunakan data asli tanpa augmentasi, sedangkan model kedua dilatih menggunakan data yang telah dioptimasi melalui proses EDA.

3.5. Evaluation

Fase kelima, *Evaluation*, bertujuan untuk mengukur tingkat keberhasilan dan kinerja kedua model tersebut menggunakan data uji. Metrik komputasional standar yang dihitung meliputi *accuracy*, *precision*, *recall*, dan *F1-score*, di mana seluruh perbandingan ketepatan prediksi divisualisasikan melalui *Confusion Matrix*. Adapun rumus matematis yang digunakan untuk menghitung setiap metrik evaluasi tersebut adalah sebagai berikut:

3.5.1 Precision

Precision merupakan metrik yang mengukur tingkat ketepatan prediksi positif yang dihasilkan model. Precision dihitung berdasarkan proporsi True Positive (TP) terhadap jumlah keseluruhan prediksi positif, termasuk False Positive (FP). Nilai precision yang tinggi menunjukkan bahwa model mampu melakukan deteksi objek secara tepat dengan jumlah kesalahan deteksi (FP) yang rendah.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

3.5.2 Recall

Recall mengukur sejauh mana model mampu mendeteksi seluruh objek yang seharusnya terdeteksi. Recall merupakan rasio antara True Positive (TP) dan jumlah total objek aktual yang terdiri dari True Positive dan False Negative (FN). Nilai recall yang tinggi menunjukkan bahwa model memiliki kemampuan deteksi yang baik dan mampu mengurangi risiko objek terlewatkan.

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

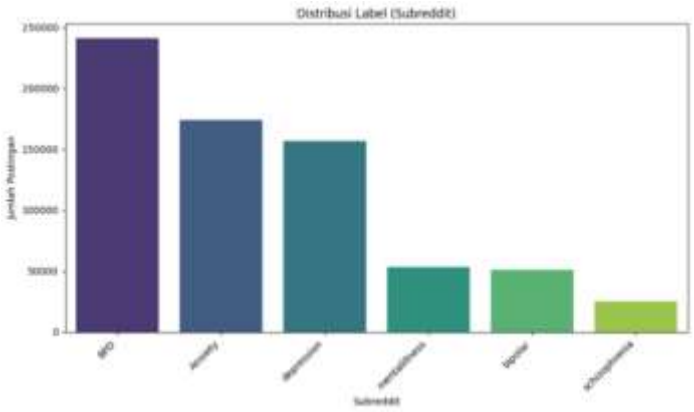
3.5.3 Accuracy (Akurasi)

Accuracy (akurasi) merupakan metrik evaluasi yang digunakan untuk mengukur tingkat ketepatan model klasifikasi secara keseluruhan, yaitu seberapa besar proporsi prediksi yang dilakukan oleh model sesuai dengan label sebenarnya. Dengan kata lain, accuracy menunjukkan persentase data yang berhasil diklasifikasikan dengan benar oleh model terhadap seluruh data pengujian.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

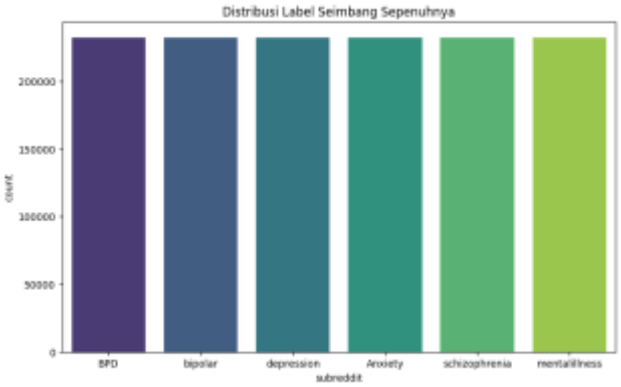
4. Hasil dan Pembahasan

Bagian tahap awal penelitian dilakukan dengan mengeksplorasi dataset dari Kaggle yang bersumber dari Reddit untuk mengidentifikasi karakteristik awal dari enam kategori gangguan kesehatan mental yang diteliti. Berdasarkan analisis statistik terhadap distribusi label, ditemukan adanya ketidakseimbangan kelas (class imbalance) yang signifikan pada data asli. Seperti yang ditunjukkan pada Gambar 2, kategori BPD memiliki jumlah sampel yang sangat dominan, sementara kategori seperti *Mentalillness* dan Schizophrenia memiliki jumlah data yang jauh lebih terbatas. Kondisi ini menuntut adanya penanganan khusus agar model tidak mengalami bias terhadap kelas mayoritas selama proses pelatihan.



Gambar 2. Distribusi Label Dataset Reddit Sebelum Augmentasi

Proses transformasi data kemudian diterapkan untuk menyiapkan dataset yang lebih berkualitas melalui serangkaian pembersihan teks. Peneliti melakukan penanganan terhadap nilai null, penghapusan karakter non-alfanumerik, serta penghapusan entri teks tidak valid seperti konten berlabel removed atau deleted. Selanjutnya, teknik Easy Data Augmentation (EDA) diimplementasikan pada data latih kelas minoritas guna menambah variasi ekspresi linguistik melalui operasi penggantian sinonim, penyisipan, pertukaran, dan penghapusan kata secara acak. Hasil dari proses augmentasi ini menciptakan distribusi label yang seimbang sepenuhnya, sehingga setiap kategori memiliki proporsi data yang setara untuk diproses oleh model klasifikasi.



Gambar 3. Distribusi Label Dataset Reddit Setelah Augmentasi

Proses transformasi data kemudian dilanjutkan dengan penerapan teknik Easy Data Augmentation (EDA) untuk memperkaya dataset, terutama pada kategori yang memiliki jumlah sampel sedikit. Peneliti menggunakan empat operasi utama EDA, yaitu penggantian sinonim (synonym replacement), penyisipan acak (random insertion), pertukaran acak (random swap), dan penghapusan acak (random deletion). Teknik ini diterapkan dengan aturan ketat agar perubahan yang dihasilkan tidak merusak makna asli dari ungkapan emosional yang ada dalam teks. Dampak dari proses augmentasi ini dapat diamati pada Gambar 3, yang menampilkan distribusi label seimbang sepenuhnya, di mana setiap kelas kini memiliki jumlah data latih yang sama besar. Setelah data siap, dilakukan ekstraksi fitur menggunakan metode TF-IDF dan dilanjutkan dengan pelatihan model menggunakan algoritma *Logistic Regression*. Hasil performa model kemudian dievaluasi secara komprehensif. Sebelum merangkum kinerja model ke dalam persentase ukur, evaluasi perbandingan prediksi divisualisasikan terlebih dahulu melalui *Confusion Matrix*. Matriks ini memetakan jumlah ketepatan tebakan model pada diagonal utama (*True Positives*) berbanding dengan kesalahan tebakan (*False Positives* dan *False Negatives*).

Gambar 4. *Confusion Matrix* model tanpa augmentasi EDA

Gambar 5. *Confusion Matrix* model tanpa augmentasi EDA

Pada *Confusion Matrix* model tanpa augmentasi EDA pada gambar 4 , terlihat bahwa sistem sangat rentan mengalami bias terhadap kelas mayoritas. Hal ini dibuktikan dengan tingginya angka *False Negative* pada kelas *Mental illness*, di mana banyak teks dari kelas tersebut justru disalahartikan oleh sistem ke dalam kelas *Depression* atau *Anxiety*. Sebaliknya, pada *Confusion Matrix* model dengan augmentasi EDA pada Gambar 5, terlihat adanya peningkatan konsentrasi nilai tebakan benar pada diagonal utama untuk kelas *Mental illness*. Pola matriks ini menegaskan bahwa penambahan variasi data latih (EDA) berhasil membuat model menjadi lebih sensitif sehingga sukses menekan angka *False Negative*. Rekapitulasi perbandingan metrik evaluasi yang diturunkan dari kedua *Confusion Matrix* tersebut kemudian disajikan secara rinci pada Tabel 1.

Table 1. Perbandingan model

	Class	**Precision**	**Recall**	**F1-score**	**Accurasy**
Tanpa Augmentasi EDA	Anxiety	0.83	0.85	0.84	0.76
	BPD	0.79	0.84	0.82	
	Bipolar	0.76	0.57	0.65	
	Depression	0.67	0.78	0.72	
	Mentalillness	0.53	0.28	0.36	
	Schizophrenia	0.73	0.52	0.61	
Menggunakan Augmentasi EDA	Anxiety	0.81	0.81	0.81	0.74
	Bpd	0.78	0.73	0.75	
	Bipolar	0.80	0.74	0.77	
	Depression	0.64	0.75	0.69	
	Mentalillness	0.62	0.56	0.59	
	Schizophrenia	0.80	0.85	0.82	

Berdasarkan data pada Tabel 1, terlihat adanya perbedaan karakteristik performa antara kedua skenario pemodelan yang diuji. Model *baseline* tanpa teknik EDA mencatatkan akurasi global yang lebih tinggi sebesar 0.76 (76%) dibandingkan model dengan EDA yang mencapai 0.74 (74%). Penurunan akurasi ini merupakan dampak dari masuknya variasi teks buatan yang bersifat *noise* pada data latih, sehingga sedikit memengaruhi stabilitas prediksi pada kelas mayoritas yang sudah dominan. Namun, jika ditinjau dari tujuan utama penelitian sebagai instrumen

penyaringan (*screening*), model yang didukung teknik EDA memberikan hasil yang jauh lebih optimal secara praktis. Hal ini dibuktikan dengan lonjakan nilai *Recall* pada kelas minoritas *Mental illness yang meningkat secara* signifikan dari 0.28 menjadi 0.56, serta kenaikan *F1-Score* dari 0.36 menjadi 0.59. Dalam konteks medis dan deteksi dini, kemampuan model untuk meminimalkan kesalahan deteksi pada kelas yang jarang muncul (*False Negative*) jauh lebih diprioritaskan daripada sekadar mengejar akurasi keseluruhan. Temuan ini mengonfirmasi bahwa penerapan augmentasi EDA berhasil memperkaya pola linguistik model sehingga menjadi lebih peka dalam mendeteksi indikasi risiko gangguan kesehatan mental. Validitas hasil ini diperkuat dengan adanya proses validasi pakar (*expert judgment*) oleh psikolog profesional. Hasil konfirmasi menunjukkan bahwa pola-pola bahasa yang berhasil diklasifikasikan oleh model NLP, terutama pada kelas-kelas yang telah diaugmentasi, memiliki relevansi klinis yang kuat dengan ungkapan keluhan pasien secara riil di lapangan. Dengan demikian, model terbaik ini memiliki landasan yang kuat untuk diimplementasikan ke dalam sistem deteksi dini mandiri berbasis web sebagai bantuan teknologi awal bagi masyarakat.

5. Kesimpulan

Penelitian ini berhasil mengimplementasikan alur kerja klasifikasi teks untuk deteksi dini indikasi gangguan kesehatan mental menggunakan kerangka kerja CRISP-DM. Melalui serangkaian proses pembersihan data dan praproses pada dataset Reddit, teks berhasil ditransformasikan menjadi data terstruktur yang siap diolah oleh model *Machine Learning*. Penerapan teknik *Easy Data Augmentation* (EDA) terbukti sangat efektif dalam mengatasi masalah ketidakseimbangan kelas (*class imbalance*) dengan meningkatkan variasi linguistik pada kategori minoritas. Meskipun penggunaan teknik EDA menyebabkan sedikit penurunan pada akurasi global dari 0.76 menjadi 0.74, model ini memberikan performa yang jauh lebih unggul dalam aspek sensitivitas deteksi. Peningkatan metrik *Recall* pada kelas *Mentalillness* yang melonjak dari 0.28 menjadi 0.56 menunjukkan bahwa model menjadi lebih peka dalam mengenali sinyal risiko yang jarang muncul. Oleh karena itu, model yang didukung oleh teknik EDA jauh lebih direkomendasikan untuk digunakan sebagai instrumen penyaringan awal (*screening*). Sistem ini telah divalidasi oleh pakar psikolog profesional dan diimplementasikan ke dalam aplikasi web sederhana sebagai bukti keberhasilan *deployment* model di lingkungan yang nyata. Berdasarkan temuan dan keterbatasan dalam penelitian ini, terdapat beberapa saran yang diusulkan untuk pengembangan selanjutnya. Pertama, penelitian mendatang diharapkan dapat menggunakan dataset asli yang berbahasa Indonesia agar hasil deteksi lebih akurat dalam menangkap konteks budaya dan ragam bahasa informal masyarakat lokal. Kedua, eksplorasi terhadap algoritma *Deep Learning* yang lebih kompleks, seperti *Bidirectional Long Short-Term Memory* (Bi-LSTM) atau model berbasis *Transformer* (seperti IndoBERT), sangat disarankan untuk meningkatkan kemampuan pemahaman konteks kalimat yang lebih dalam. Terakhir, pengembangan sistem aplikasi di masa depan sebaiknya diintegrasikan secara langsung dengan fasilitas layanan kesehatan mental profesional agar setiap indikasi risiko yang terdeteksi dapat segera mendapatkan validasi dan penanganan medis yang tepat.

Daftar pustaka

[1] B. A. Mustofa, W. Laksito, and Y. Saptomo, "Use of Natural Language Processing in Social Media Text Analysis," vol. 4, no. 2, pp. 2–5, 2025, doi: https://doi.org/10.59934/jaiea.

[2] T. Zhang, "Natural language processing applied to mental illness detection : a narrative review," no. December, pp. 1–13, 2021, doi: 10.1038/s41746-022-00589-7.

[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," vol. 6, no. 2, pp. 557–574, 2025, doi: DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4206.

[4] A. Salsabila, S. Fitriyani, V. Rahmayanti, and S. Nastiti, "Pendekatan Linguistik dalam Klasifikasi Emosi Depresi untuk Deteksi Dini Kesehatan Mental di Reddit," pp. 771–781, 2025, doi: 10.33364/algoritma/v.22-2.2927.

[5] M. Mohan, T. Ayyalasomayajula, A. Agarwal, and S. Khan, "Reddit social media text analysis for depression prediction : using logistic regression with enhanced term frequency-inverse document frequency features," vol. 14, no. 5, pp. 5998–6005, 2024, doi: 10.11591/ijece.v14i5.pp5998-6005.

[6] S. Inamdar, R. Chapekar, S. Gite, and B. Pradhan, "Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing," *Human-Centric Intell. Syst.*, vol. 3, no. 2, pp. 80–91, 2023, doi: 10.1007/s44230-023-00020-8.

[7] J. Sibarani, R. Manalu, D. P. Hutasoit, W. A. Telaumbanua, and V. A. E. P, "Deteksi Kecenderungan Kesehatan Mental Menggunakan Naïve Bayes Berdasarkan Aktivitas Media Sosial Detection of Mental Health Tendencies Using Naïve Bayes Based on Social Media Activity," vol. 4, no. 2, pp. 80–87, 2025, doi: 10.55123/jomlai.v4i2.5959.

[8] A. N. Azizah *et al.*, "EASY DATA AUGMENTATION UNTUK DATA YANG IMBALANCE PADA EASY DATA AUGMENTATION FOR IMBALANCED DATA OF ONLINE HEALTH," vol. 10, no. 5, 2023, doi: 10.25126/jtiik.2023107082.

[9] M. Zidan, A. Pinandito, and A. N. Rusydi, "Analisis Perbandingan Efisiensi Metode Augmentasi Data EDA , Back- Translation Dan LLM Untuk Klasifikasi Topik Berita Berbahasa Indonesia Pada Kondisi Class Imbalance," vol. 10, no. 1, pp. 1–6, 2026, doi: DOI: 10.59095/ijcsr.v10i1.1590.

[10] E. E. Pratiwi, A. R. Aisy, N. Ananta, and T. Informatika, "KLASIFIKASI KESEHATAN MENTAL PADA," vol. 13, no. 2, 2025.

[11] B. Jacennik, E. Zawadzka-gosk, and J. P. Moreira, "Evaluating Patients ' Experiences with Healthcare Services : Extracting Domain and Language-Specific Information from Free-Text Narratives," 2022, doi: DOI: 10.1038/s41598-022-12345-x.

[12] N. Surojudin, S. Butsianto, and A. Firmansyah, "BULLETIN OF COMPUTER SCIENCE RESEARCH Perbandingan Kinerja Naïve Bayes dengan dan Tanpa SMOTE untuk Klasifikasi Gangguan Kecemasan Mahasiswa pada Data Tidak Seimbang," vol. 6, no. 2,

pp. 804–812, 2026, doi: 10.47065/bulletincsr.v6i2.1021.
[13] P. Guleria, "NLP based text classification using TF-IDF enabled fine-tuned long short-term memory : An empirical analysis," *Array*, vol. 27, no. August, p. 100467, 2025, doi: 10.1016/j.array.2025.100467.
[14] J. A. Adeyiga, "Fake News Detection Using a Logistic Regression Model and Natural Language Processing Techniques," pp. 1–18, 2023, doi: DOI: 10.48550/arXiv.2305.12345.
[15] N. Amin, "Comparative Evaluation of Logistic Regression and Naïve Bayes for Fake News Detection Using NLP Techniques Comparative Evaluation of Logistic Regression and Naïve Bayes for Fake News Detection Using NLP Techniques," pp. 0–15, 2025, doi: 10.20944/preprints202512.0630.v1.
[16] H. Utama, E. Daniati, and A. Masruro, "WEAK SUPERVISION DENGAN PENDEKATAN LABELING FUNCTION UNTUK ANALISIS SENTIMEN," vol. 3, no. 1, pp. 49–57, 2024, doi: DOI: 10.59095/ijcsr.v3i1.93.
[17] F. F. Pradana, K. Auliasari, and M. Orisa, "ANALISIS SENTIMEN TERHADAP GENERASI Z DALAM ETIKA KERJA PADA MEDIA SOCIAL MENGGUNAKAN METODE SUPPORT VECTOR MACHINE," vol. 9, no. 6, pp. 10675–10682, 2025, doi: DOI: https://doi.org/10.36040/jati.v9i6.15463.
[18] K. Rahayu, V. Fitria, and D. Septhya, "Text Classification for Detecting Depression and Anxiety among Twitter Users based on Machine Learning Klasifikasi Teks untuk Mendeteksi Depresi dan Kecemasan Pada Pengguna Twitter Berbasis Machine Learning," vol. 3, no. October, pp. 108–114, 2023.
[19] P. C. Bankruptcy, "Penerapan Metode Oversampling SMOTE Pada Algoritma Random Forest Untuk Prediksi Kebangkrutan Perusahaan," vol. 22, no. 1, pp. 207–214, 2023, doi: 10.33633/tc.v22i1.7527.
[20] S. Method, "Deteksi Dini Gangguan Kesehatan Mental berdasarkan Sentimen menggunakan Metode Stacking Early Detection of Mental Health Disorders based on Sentiment using," vol. 14, pp. 271–280, 2025, doi: 10.29103/sistemasi.v14i1.1396.
[21] W. Nursahid, B. I. Nugroho, and S. Syefudin, "Optimalisasi Preprocessing Data Menggunakan Pendekatan CRISP-DM untuk Meningkatkan Kualitas Klasifikasi Penyakit Jantung," vol. 4, no. 3, pp. 3621–3626, 2025, doi: 10.48550/arXiv.2301.04521.