

Penerapan Natural Language Processing dan Machine Learning untuk Prediksi Stres Siswa SMA Berdasarkan Analisis Teks

Muhammad Rofiq Sudrajat*, Muhammad Zakariyah

Sains dan Teknologi, Informatika, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia

Email: ^{1,*}moh.rofiqs123@gmail.com, ²muhammad.zakariyah@staff.uty.ac.id

Email Penulis Korespondensi: moh.rofiqs123@gmail.com

Submitted: 31/10/2024; Accepted: 01/12/2024; Published: 03/12/2024

Abstrak—Penelitian ini mengeksplorasi penerapan Natural Language Processing (NLP) dan Machine Learning dalam memprediksi stres di kalangan siswa Sekolah Menengah Atas (SMA). Karena stres pada siswa sering kali tidak disadari, maka diperlukan metode yang efektif untuk mengidentifikasi lebih dini. Untuk mengatasi masalah ini, penelitian ini mengembangkan model prediksi stres berbasis teks menggunakan NLP untuk ekstraksi fitur dan Machine Learning untuk klasifikasi. Teknik inti NLP meliputi pembersihan data, penghapusan stopword, tokenisasi, dan lemmatisasi untuk memproses data teks, sedangkan ekstraksi fitur dicapai melalui metode seperti Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), dan N-grams (Unigram, Bigram, Trigram). Model Machine Learning yang diujicobakan antara lain Logistic Regression, Naive Bayes, Random Forest, dan Support Vector Machine (SVM). Hasil dari eksperimen menunjukkan bahwa model Naive Bayes yang menggunakan fitur Bigram mencapai akurasi tertinggi yaitu 95,6%, dengan model lainnya mencapai sekitar 93%. Terlepas dari kinerja model yang kuat, kesalahan seperti False Positive dan False Negative masih ditemukan, yang mengindikasikan adanya ruang untuk perbaikan. Penelitian ini menunjukkan bahwa NLP yang dikombinasikan dengan Machine Learning memberikan pendekatan yang efektif untuk mengidentifikasi stres siswa, dengan potensi yang menjanjikan untuk intervensi kesehatan mental di lingkungan pendidikan.

Kata Kunci: Natural Language Processing; Machine Learning; Prediksi Stres; Siswa SMA; Analisis Teks

Abstract—This research explores the application of Natural Language Processing (NLP) and Machine Learning in predicting stress among high school students. As stress in students often goes unnoticed, there is a need for effective methods to identify it early. To address this issue, this research develops a text-based stress prediction model using NLP for feature extraction and Machine Learning for classification. Core NLP techniques include data cleaning, stopword removal, tokenization, and lemmatization to process text data, while feature extraction is achieved through methods such as Bag Of Words (BOW), Term Frequency-Inverse Document Frequency (TF-IDF), and N-grams (Unigram, Bigram, Trigram). The Machine Learning models tested include Logistic Regression, Naive Bayes, Random Forest, and Support Vector Machine (SVM). Results from the experiments showed that the Naive Bayes model using Bigram features achieved the highest accuracy of 95.6%, with the other models achieving around 93%. Despite the strong performance of the models, errors such as False Positive and False Negative were still found, indicating room for improvement. This research shows that NLP combined with Machine Learning provides an effective approach to identifying student stress, with promising potential for mental health interventions in educational settings.

Keywords: Natural Language Processing, Machine Learning, Stress Prediction, High School Students, Text Analysis

1. PENDAHULUAN

Kesehatan mental merupakan hal yang sangat penting bagi remaja, terutama siswa di Sekolah Menengah Atas (SMA), karena mereka mengalami banyak perubahan fisik, emosional dan sosial yang dapat mempengaruhi kesehatan psikologis mereka [1]. Salah satu masalah yang paling umum adalah stres, stres dapat berdampak negatif pada kesehatan mental, prestasi akademik, hubungan sosial dan pertumbuhan emosional [2]. Data di seluruh dunia menunjukkan bahwa 5% hingga 70% remaja mengalami stres dan kecemasan [3]. Di Indonesia, sekitar 14 juta orang mengalami gangguan emosional dan stres ringan serta 6% remaja di atas 15 tahun mengalami stres [4]. Siswa di SMA menghadapi banyak tantangan, termasuk tuntutan akademik, persaingan dengan teman sebaya dan tekanan untuk mencapai tujuan mereka, yang dapat menyebabkan stres [5]. Kondisi stres yang berlebihan dapat menyebabkan masalah kesehatan mental seperti kecemasan, depresi dan masalah perilaku [6]. Oleh karena itu, penting untuk mengidentifikasi dan menangani stres secara dini agar siswa dapat menjalani masa remaja mereka dengan lebih sehat dan produktif.

Belakangan ini, penggunaan teknologi *Artificial Intelligence* (AI), khususnya *Natural Language Processing* (NLP), telah menjadi metode yang menarik untuk mengidentifikasi masalah kesehatan mental melalui analisis teks [7]. NLP adalah ilmu kecerdasan buatan yang memungkinkan komputer untuk memahami, menafsirkan dan menghasilkan bahasa manusia [8]. Para peneliti telah menunjukkan ketertarikan yang semakin besar dalam menggunakan NLP untuk prediksi stres karena kemampuannya untuk mengidentifikasi pola-pola bahasa yang terkait dengan kondisi psikologis [9], [10], [11], [12]. Dengan memanfaatkan teknologi ini, diharapkan bisa untuk mendeteksi tanda-tanda awal stres yang mungkin terabaikan dalam interaksi sehari-hari.

Dalam penelitian ini, beberapa algoritma *Machine Learning* akan diterapkan untuk memprediksi stres berdasarkan data yang telah dianalisis menggunakan NLP. Algoritma seperti *Logistic Regression*, *Naive Bayes*, *Random Forest*, dan *Support Vector Machine* (SVM) telah terbukti efektif dalam mengklasifikasikan dan memprediksi berbagai kategori teks berdasarkan fitur-fitur yang diekstrak dari data [13], [14], [15]. Dengan melatih

model-model ini menggunakan dataset yang beragam, diharapkan dapat diperoleh akurasi tinggi dalam mendeteksi stres siswa berdasarkan pola bahasa yang teridentifikasi.

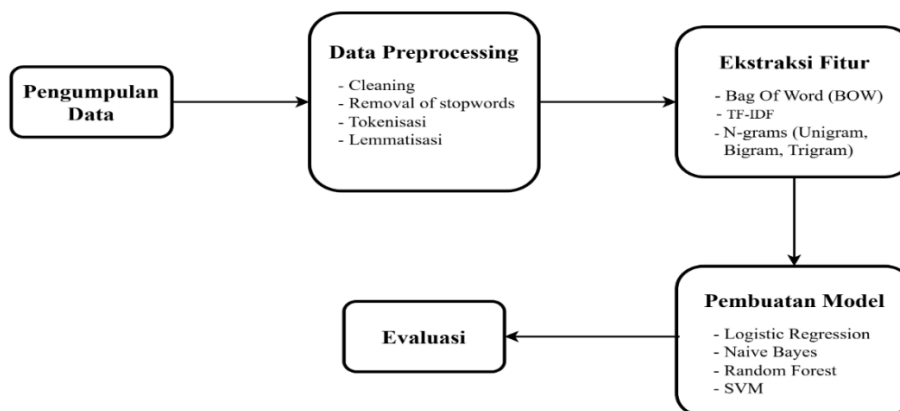
NLP dan algoritma *Machine Learning* semakin banyak digunakan untuk mendeteksi masalah kesehatan mental melalui analisis teks. Penelitian Low dkk. menunjukkan bahwa NLP dapat membantu menganalisis perilaku terkait kecemasan selama pandemi *COVID-19*, dengan menemukan pola bahasa yang berkaitan dengan kecemasan kesehatan dalam forum dukungan reddit [16]. Sementara itu, Najamuddin dkk. mengimplementasikan *chatbot* untuk deteksi dini depresi pada mahasiswa menggunakan *Patient Health Questionnaire* (PHQ-9) dan NLP. Hasilnya menunjukkan bahwa sekitar 84% partisipan merasa bahwa *chatbot* ini efektif dalam membantu deteksi dini depresi [17]. Studi oleh Kim dkk. memanfaatkan *Logistic Regression* untuk mengklasifikasikan teks terkait gejala depresi berdasarkan data dari media sosial, di mana model yang dikembangkan mampu mengidentifikasi kalimat yang menunjukkan gejala depresi berdasarkan kuesioner PHQ-9 dan mencapai akurasi yang lebih baik dibandingkan model lainnya, sehingga menekankan pentingnya data teks dalam penilaian kesehatan mental [18]. Penelitian oleh Syahputra dkk. menerapkan algoritma Naive Bayes untuk menganalisis sentimen terkait kesehatan mental di kalangan remaja melalui data Twitter, mengumpulkan data dengan hashtag kesehatan mental dan mengklasifikasikan sentimen menjadi positif, negatif, atau netral dengan tingkat akurasi mencapai 80%, memberikan wawasan tentang persepsi masyarakat terhadap isu kesehatan mental [19]. Selain itu, studi oleh Bader dkk. mengeksplorasi penggunaan *Random Forest* dalam mendeteksi tingkat keparahan depresi berdasarkan analisis teks dari kuesioner kesehatan mental, dan penelitian ini menunjukkan bahwa *Random Forest* dapat mencapai *Area Under the Curve* (AUC) hingga 0,84 pada data dari 830 peserta [20]. Selanjutnya, Fadhila dkk. mengevaluasi hubungan antara kualitas tidur dengan stres menggunakan algoritma *K-Nearest Neighbor* (KNN), *Decision Tree*, dan SVM. Hasil menunjukkan bahwa algoritma KNN dan SVM masing-masing mencapai akurasi sempurna 100%, sedangkan *Decision Tree* mencapai akurasi 99%, yang menekankan pentingnya analisis fitur seperti pola tidur dalam deteksi dini stres [21].

Kesimpulan dari tinjauan pustaka menunjukkan bahwa penerapan NLP dan *Machine Learning* dalam analisis stres siswa SMA memiliki potensi yang besar. Berbagai penelitian telah mengonfirmasi efektivitas teknik NLP dalam menganalisis teks untuk deteksi awal masalah kesehatan mental, seperti kecemasan, stres, dan depresi, serta bagaimana algoritma *Machine Learning* seperti *Logistic Regression*, *Naive Bayes*, *Random Forest*, dan *Support Vector Machine* (SVM) dapat digunakan untuk klasifikasi stres dengan akurasi yang bervariasi.

Tujuan penelitian ini adalah untuk mengeksplorasi potensi penggunaan NLP dan *Machine Learning* untuk memprediksi stres siswa dengan memanfaatkan berbagai teknik ekstraksi fitur. Dengan menggunakan berbagai ekstraksi fitur seperti *Bag of Words* (BoW), *Term Frequency-Inverse Document Frequency* (TF-IDF), dan *N-grams* (*Unigram*, *Bigram*, *Trigram*), penelitian ini bertujuan untuk menemukan metode atau model yang paling efektif dalam memprediksi stres melalui analisis teks. Hasil dari penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai kesehatan mental siswa serta membuka peluang untuk pengembangan sistem prediksi yang lebih efisien dan efektif dalam mendukung intervensi kesehatan mental.

2. METODOLOGI PENELITIAN

Metodologi penelitian dalam penelitian ini disusun dalam beberapa tahap sistematis yang meliputi pengumpulan data, *preprocessing*, ekstraksi fitur, pembuatan model dan evaluasi. Tahapan ini dirancang untuk memastikan bahwa penelitian ini dapat menghasilkan hasil yang akurat dan andal dalam memprediksi stres pada siswa SMA. Berikut adalah tahapan dalam penelitian yang dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1 Pengumpulan Data

Proses pengumpulan data dalam penelitian ini melibatkan dua tahap utama yaitu mengembangkan dataset teks berdasarkan indikator terkait stres dan memvalidasinya melalui wawancara dengan para ahli psikologi. Dataset yang

berisi 1.476 data teks ini dibuat dengan menggunakan indikator dan kata kunci dari sumber-sumber terkemuka di bidang psikologi dan kesehatan mental. Ini termasuk "kecemasan", "ketegangan", "kelelahan", "kecewa", "kewalahan", "kehilangan kendali", "depresi" dan "panik" [22], [23], [24], [25], [26].

Untuk memastikan relevansi dan keakuratan dataset, wawancara mendalam dilakukan dengan psikolog dan konselor berpengalaman yang memahami tentang stres pada remaja, khususnya dalam konteks pendidikan. Dataset secara manual diberi label 1 = "stres" atau 0 = "tidak stres" mengikuti pendekatan dari dataset Dreddit yang digunakan untuk analisis stres di media sosial oleh Kasmin dkk [27]. Seperti terlihat pada Tabel 1 berikut, kombinasi antara dataset yang dikembangkan berdasarkan sumber terpercaya dan validasi dari ahli psikologi memberikan dasar yang kuat untuk pengembangan model prediksi stres yang akurat dan kontekstual.

Tabel 1. Data teks

Teks	Label
"Saya merasa kelelahan ketika tidak ada waktu untuk bersantai dan bersenang-senang."	1
"Kelelahan fisik dan mental sering kali membuat saya merasa terasing."	1
"Saya merasa lelah ketika menghadapi konflik yang berkepanjangan."	1
"Kelelahan ini membuat saya merasa cemas akan tugas yang belum selesai."	1
"Saya merasa sangat lelah setelah menjalani kegiatan yang sangat padat."	1
"Saya merasa segar dan bertenaga setelah tidur yang nyenyak."	0
"Tidak merasa lelah membuat saya lebih produktif saat belajar."	0
"Saya merasa energik setelah berolahraga di pagi hari."	0
"Saya bisa fokus dan tidak merasa lelah saat belajar bersama teman."	0
"Beristirahat sejenak membuat saya merasa lebih baik dan tidak lelah."	0

2.2 Data Preprocessing

Pada tahap ini data teks diproses terlebih dahulu menggunakan beberapa teknik NLP untuk membersihkan dan mempersiapkannya untuk dianalisis. Data mentah diubah menjadi data yang lebih terstruktur yang dapat dianalisis. Beberapa langkah yang dilakukan dalam proses ini meliputi:

a. Cleaning

Pada tahap ini dilakukan penghapusan karakter-karakter khusus seperti tanda baca, simbol, dan angka yang tidak relevan untuk analisis teks. Selain itu, semua kata dalam dataset diubah menjadi huruf kecil (*lowercasing*) untuk memastikan konsistensi dalam analisis.

b. Hapus Stopwords

Langkah ini melibatkan penghapusan kata-kata umum yang tidak memberikan kontribusi signifikan terhadap pemahaman makna teks dalam konteks analisis stres. Kata-kata seperti "dan", "atau", "sebuah" dan "yang" sering kali tidak mengandung informasi penting dalam analisis teks, sehingga mereka dihapus untuk memfokuskan perhatian pada kata-kata yang lebih relevan.

c. Tokenisasi

Tokenisasi adalah proses pemisahan teks menjadi unit-unit yang lebih kecil, yang disebut token. Token ini bisa berupa kata, frasa, atau kalimat, tergantung pada tujuan analisis. Tokenisasi membantu dalam mengidentifikasi elemen-elemen penting dalam teks yang dapat digunakan dalam pengolahan lebih lanjut.

d. Lematisasi

Pada tahap lemmatisasi, kata-kata diubah menjadi bentuk dasarnya untuk menyamakan makna kata yang berbeda bentuk. Misalnya, kata "berlari", "berlari-lari", dan "berlarian" akan diubah menjadi bentuk dasar "lari". Langkah ini penting untuk mengurangi kompleksitas dan memastikan bahwa variasi kata yang memiliki makna sama diperlakukan sebagai satu entitas.

2.3 Ekstraksi Fitur

Ekstraksi fitur melibatkan perubahan teks ke dalam format numerik agar dapat ditafsirkan oleh algoritma *Machine Learning*. Tujuan dari proses ini adalah untuk menangkap karakteristik penting dalam teks yang dapat membantu dalam memprediksi stres. Penelitian ini menggunakan beberapa teknik ekstraksi fitur, termasuk:

a. Bag of Words (BoW)

Metode BoW adalah teknik yang mengubah teks menjadi representasi berdasarkan frekuensi kemunculan kata-kata, tanpa mempertimbangkan urutan atau konteks kata tersebut. Dalam BoW, setiap kata dianggap sebagai fitur yang terpisah, dan vektor fitur yang dihasilkan menunjukkan berapa kali setiap kata muncul dalam dokumen. BoW membantu model dalam mengidentifikasi kata-kata yang sering muncul dalam teks yang berhubungan dengan stres, seperti "kecemasan" atau "beban", yang bisa menjadi indikator penting dalam klasifikasi stres atau tidak stres..

b. Term Frequency-Inverse Document Frequency (TF-IDF)

Berbeda dengan pendekatan BoW, metode TF-IDF menilai tingkat kepentingan sebuah kata dalam kumpulan dokumen yang lebih besar. TF-IDF memperhitungkan frekuensi kemunculan kata dalam dokumen tertentu (*Term Frequency*) dan distribusinya di seluruh kumpulan dokumen (*Inverse Document Frequency*). Kata-kata yang

sering muncul di satu dokumen tetapi jarang muncul di dokumen lain akan mendapatkan skor TF-IDF yang lebih tinggi. Teknik ini menyoroti kata-kata yang lebih spesifik dan relevan untuk mengidentifikasi stres, seperti “kewalahan” atau “panik”, daripada kata-kata yang umum.

c. N-grams (Unigram, Bigram, Trigram)

Teknik N-grams melibatkan penggabungan kata-kata secara berurutan untuk membentuk fitur yang lebih kompleks, seperti *unigram* (satu kata), *bigram* (dua kata), atau *trigram* (tiga kata). Pendekatan ini sangat berguna untuk menangkap pola-pola frasa atau ungkapan yang lebih bermakna daripada hanya kata individu. Misalnya, frasa seperti "sangat khawatir" atau "beban berat" sering kali lebih menggambarkan keadaan stres daripada hanya menggunakan kata-kata tunggal. Dengan N-grams, model dapat mengenali pola atau ekspresi yang lebih kontekstual yang sering ditemukan pada teks yang berkaitan dengan stres.

2.4 Pembuatan Model

Setelah melalui tahap *preprocessing* dan ekstraksi fitur, proses pembuatan model dimulai dengan memisahkan data ke dalam data latih dan data uji atau *train-test split* untuk setiap jenis ekstraksi fitur. Pemisahan ini membuat model hanya melatih pada sebagian data dan sebagian data lainnya digunakan untuk mengevaluasi kinerja model dengan kombinasi ekstraksi fitur dan algoritma *Machine Learning* untuk memastikan hasil prediksi yang akurat. Berbagai algoritma *Machine Learning* diterapkan untuk membuat model prediksi stres, meliputi:

a. Logistic Regression

Logistic Regression adalah algoritma klasifikasi yang banyak digunakan untuk mengerjakan soal biner, seperti menentukan apakah seseorang mengalami stres atau tidak. Algoritma ini memprediksi probabilitas suatu kelas dengan menggunakan kombinasi fitur linear. Dengan mengasumsikan hubungan linear antara variabel input dan output, *Logistic Regression* ideal untuk klasifikasi dengan kategori yang jelas, seperti “stres” dan “tidak stres”.

b. Naive Bayes

Naive Bayes adalah algoritma berbasis probabilitas yang bekerja dengan asumsi independensi fitur. Algoritma ini sangat efektif untuk memproses data teks, terutama ketika data memiliki dimensi fitur yang tinggi. *Naive Bayes* sangat cocok untuk tugas-tugas seperti analisis sentimen dan klasifikasi teks, dan dapat menghasilkan model yang cepat dan efisien, bahkan dengan asumsi independensi yang disederhanakan.

c. Random Forest

Random Forest adalah teknik ensemble yang menggabungkan beberapa pohon keputusan untuk membuat prediksi yang lebih andal dan tepat. Setiap pohon di dalam kumpulan membuat prediksi, dan hasil akhirnya ditentukan oleh suara terbanyak. Metode ini sangat efektif untuk mengelola data yang kompleks dengan fitur-fitur yang saling terkait. *Random Forest* membantu meminimalkan risiko *overfitting* dan meningkatkan akurasi, terutama dengan set data yang lebih besar.

d. Support Vector Machine (SVM)

SVM adalah algoritma klasifikasi yang bekerja dengan membagi kelas data menggunakan *hyperplane* yang optimal, memastikan batas maksimum di antara keduanya. Algoritma ini sangat efektif untuk tugas klasifikasi biner, terutama ketika data dipisahkan dengan jelas atau memiliki ruang fitur berdimensi tinggi. SVM biasanya memberikan hasil yang kuat ketika berhadapan dengan data yang memiliki banyak fitur atau menunjukkan hubungan non-linear.

2.5 Evaluasi

Dalam penelitian ini, evaluasi model dilakukan dengan menggunakan *confusion matrix* dan *classification report* untuk menganalisis kemampuan model dalam mengidentifikasi dan memprediksi stres di kalangan siswa SMA. *Confusion matrix* memberikan gambaran yang jelas mengenai prediksi yang benar dan salah untuk setiap kelas (stres dan tidak stres), membantu mengidentifikasi kesalahan spesifik, seperti *False Positive* dan *False Negative*. Di sisi lain, *classification report* menawarkan serangkaian matriks evaluasi yang lebih rinci, termasuk *accuracy*, *precision*, *recall*, dan *F1-score* untuk setiap kelas. Evaluasi model dilakukan dengan menggunakan data uji yang terpisah dari data pelatihan untuk memastikan bahwa model dapat menggeneralisasi dengan baik pada data baru yang belum pernah digunakan sebelumnya. Hasil evaluasi memberikan wawasan yang sangat berharga mengenai kekuatan dan kelemahan masing-masing model dan menyarankan untuk peningkatan dan pengoptimalan lebih lanjut.

3. HASIL DAN PEMBAHASAN

Pada bagian ini, akan disajikan hasil eksperimen dan pembahasan mengenai penggunaan algoritma *Machine Learning* untuk memprediksi stres pada siswa SMA. Penelitian ini bertujuan untuk mengeksplorasi efektivitas berbagai algoritma *Machine Learning* dalam menganalisis teks dan memprediksi stres pada siswa. Melalui serangkaian eksperimen dan pengujian model, penelitian ini berupaya untuk mengidentifikasi metode yang paling akurat dan dapat diandalkan dalam memprediksi stres, serta memberikan wawasan lebih dalam mengenai penerapan teknologi *Natural Language Processing* (NLP) dan *Machine Learning* untuk mendukung kesehatan mental di lingkungan pendidikan.

3.1 Hasil Eksperimen

Penelitian ini melibatkan serangkaian eksperimen yang bertujuan untuk menilai dan membandingkan kinerja berbagai model *Machine Learning* untuk memprediksi stres di kalangan siswa SMA. Eksperimen ini menguji kombinasi metode ekstraksi fitur (BoW, TF-IDF, *Unigram*, *Bigram* dan *Trigram*) dengan beberapa algoritma *Machine Learning*, seperti *Logistic Regression*, *Naive Bayes*, *Random Forest* dan *SVM*. Setiap model dievaluasi berdasarkan kinerja dalam memprediksi stres pada data uji yang telah dipisahkan dengan data latih sebelumnya. Dengan menguji berbagai teknik ekstraksi fitur dan algoritma *Machine Learning*, penelitian ini bertujuan untuk menemukan kombinasi yang memberikan hasil prediksi yang paling akurat atau memiliki akurasi yang tinggi dan efisien dalam konteks stres siswa SMA. Ringkasan hasil eksperimen yang diperoleh dari setiap kombinasi ekstraksi fitur dan algoritma *Machine Learning* yang digunakan dalam penelitian ini ditunjukkan pada Tabel 2.

Tabel 2. Hasil eksperimen dengan data latih

No	Model	Ekstraksi Fitur	Accuracy	
			Train	Test
1	<i>Logistic Regression</i>	BoW	98.1%	92.9%
2	<i>Logistic Regression</i>	TF-IDF	95.8%	93.4%
3	<i>Logistic Regression</i>	<i>Unigram</i>	95.8%	93.4%
4	<i>Logistic Regression</i>	<i>Bigram</i>	96.2%	94%
5	<i>Logistic Regression</i>	<i>Trigram</i>	95.3%	91.5%
6	<i>Naive Bayes</i>	BoW	96%	94.8%
7	<i>Naive Bayes</i>	TF-IDF	96.7%	94.3%
8	<i>Naive Bayes</i>	<i>Unigram</i>	96.7%	94.3%
9	<i>Naive Bayes</i>	<i>Bigram</i>	96.7%	95.6%
10	<i>Naive Bayes</i>	<i>Trigram</i>	96.2%	93.2%
11	<i>Random Forest</i>	BoW	100%	93.4%
12	<i>Random Forest</i>	TF-IDF	100%	92.1%
13	<i>Random Forest</i>	<i>Unigram</i>	100%	92.1%
14	<i>Random Forest</i>	<i>Bigram</i>	100%	93.2%
15	<i>Random Forest</i>	<i>Trigram</i>	100%	92.9%
16	<i>SVM</i>	BoW	98.5%	92.9%
17	<i>SVM</i>	TF-IDF	97.7%	94.5%
18	<i>SVM</i>	<i>Unigram</i>	97.7%	94.5%
19	<i>SVM</i>	<i>Bigram</i>	97.1%	94.3%
20	<i>SVM</i>	<i>Trigram</i>	97.3%	93.7%

Dari 20 eksperimen yang dilakukan dengan data latih, terdapat 4 model dengan akurasi terbaik yang muncul dari berbagai algoritma dan ekstraksi fitur, yang dapat dilihat pada Tabel 3.

Tabel 3. Model terbaik dengan data latih

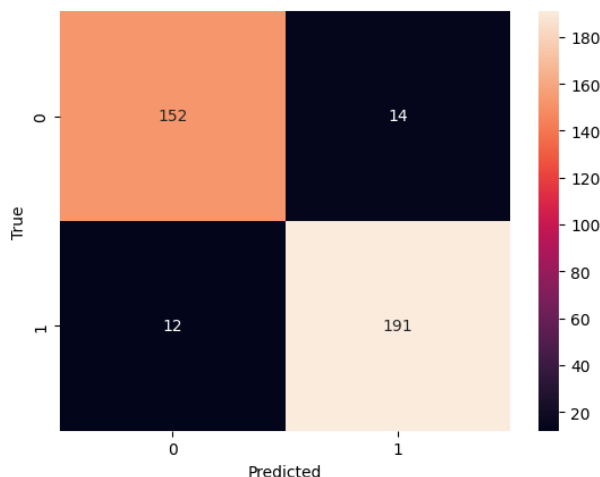
No	Model	Ekstraksi Fitur	Accuracy	
			Train	Test
1	<i>Logistic Regression</i>	BoW	98.1%	92.9%
2	<i>Naive Bayes</i>	<i>Bigram</i>	96.7%	95.6%
3	<i>Random Forest</i>	BoW	100%	93.4%
4	<i>SVM</i>	BoW	98.5%	92.9%

Hasil eksperimen menunjukkan bahwa model *Random Forest* memiliki akurasi tertinggi, mencapai 100% pada dataset pelatihan untuk semua ekstraksi fitur yang digunakan (BoW, TF-IDF, *Unigram*, *Bigram*, dan *Trigram*). Namun, rendahnya akurasi pada data uji mengindikasikan adanya *overfitting*, sehingga performa model ini kurang optimal untuk prediksi pada data baru. *SVM* dan *Logistic Regression* juga memberikan akurasi pelatihan tinggi dengan nilai di atas 95%, serta menunjukkan perbedaan minimal antara akurasi training dan testing menandakan kemampuan generalisasi yang cukup baik.

Model dengan performa terbaik secara keseluruhan dalam eksperimen ini adalah *Naive Bayes* dengan metode *Bigram*, yang menghasilkan akurasi 95,6% pada dataset pengujian. Model ini mampu menjaga keseimbangan antara akurasi *training* dan *testing*, menunjukkan generalisasi yang optimal. Secara umum, semua model menunjukkan performa yang cukup memuaskan dengan akurasi testing di atas 90%. Selain itu, perbedaan antara akurasi *training* dan *testing* yang tidak terlalu besar (kecuali pada *Random Forest*) menunjukkan bahwa sebagian besar model memiliki generalisasi yang baik dan mampu menangani data baru dengan cukup efektif. Hasil ini menunjukkan bahwa pendekatan NLP untuk ekstraksi fitur dan *Machine Learning* yang digunakan cukup efektif dalam mengidentifikasi dan memprediksi stres pada siswa SMA.

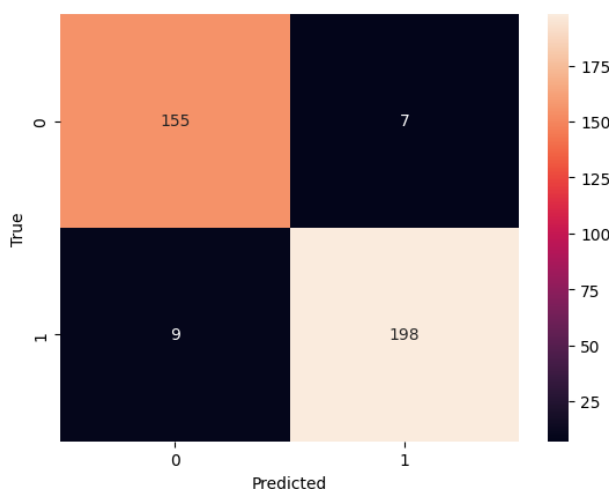
3.2 Pengujian Model

Tujuan dari pengujian adalah untuk mengevaluasi kemampuan model-model terbaik dari semua eksperimen dalam memprediksi stres pada siswa SMA. Dengan menggunakan dataset yang telah diproses dan dibagi menjadi data pelatihan dan pengujian, tahap ini berfokus pada seberapa baik model dapat mengidentifikasi dan memprediksi stres berdasarkan analisis teks. Pengujian model dilakukan menggunakan data uji yang terpisah dari data latih, guna mengukur keakuratan prediksi model. Pengukuran yang digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-score*, sehingga memberikan gambaran yang lebih komprehensif tentang kinerja masing-masing model.



Gambar 2. *Confusion matrix* model terbaik dari *Logistic Regression*

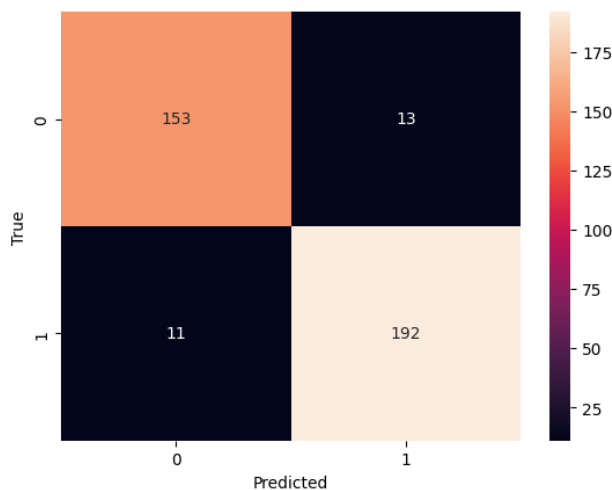
Berdasarkan *confusion matrix* yang terlihat pada Gambar 2 di atas, model *Logistic Regression* dengan metode BoW menunjukkan performa yang cukup baik dalam memprediksi stres siswa. Model ini berhasil melakukan 343 prediksi benar dari total 369 data, dengan 152 siswa yang benar-benar tidak stres teridentifikasi dengan tepat (*True Negative*) dan 191 siswa yang stres diprediksi dengan benar (*True Positive*). Namun, terdapat 14 kesalahan *False Positive*, di mana siswa yang seharusnya tidak stres diprediksi sebagai stres dan 12 kesalahan *False Negative*, di mana siswa yang sebenarnya stres tidak teridentifikasi oleh model. Kesalahan *False Negative* ini lebih kritis karena dapat mengakibatkan siswa yang membutuhkan intervensi tidak mendapat perhatian. Secara keseluruhan, model *Logistic Regression* dengan BoW memiliki kemampuan prediksi yang baik, tetapi peningkatan lebih lanjut diperlukan untuk meminimalkan kesalahan prediksi, terutama pada kasus siswa yang benar-benar stres agar tidak terlewatkan.



Gambar 3. *Confusion matrix* model terbaik dari *Naive Bayes*

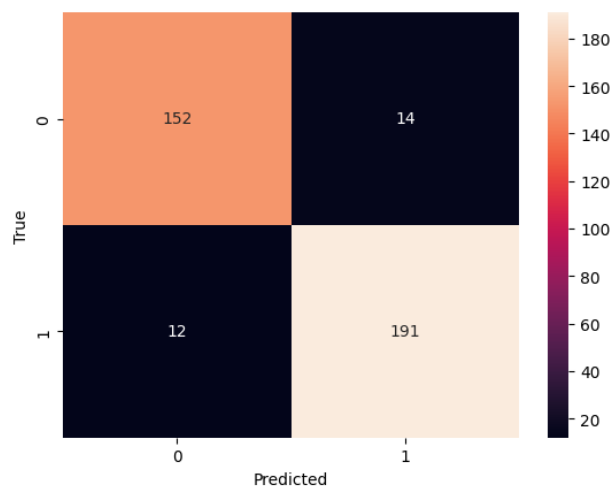
Pada Gambar 3 di atas, model *Naive Bayes* dengan metode *Bigram* menunjukkan performa yang sangat baik dalam memprediksi stres siswa. Dari total 369 data, model ini berhasil melakukan 353 prediksi benar dengan 155 siswa yang tidak stres diprediksi dengan benar (*True Negative*) dan 198 siswa yang stres juga teridentifikasi dengan tepat (*True Positive*). Namun, terdapat beberapa kesalahan prediksi, yaitu 7 kasus *False Positive* di mana siswa yang sebenarnya tidak stres diprediksi sebagai stres dan 9 kasus *False Negative* di mana siswa yang mengalami stres tidak terdeteksi oleh model. Meski terdapat kesalahan tersebut, jumlahnya relatif kecil dibandingkan total data, menunjukkan bahwa model ini memiliki kemampuan prediksi yang sangat baik. Secara keseluruhan, *Naive Bayes*

dengan *Bigram* terbukti efektif untuk mendeteksi stres siswa, dengan minimnya kesalahan prediksi. Meskipun demikian, perhatian tetap diperlukan untuk mengurangi *False Negative* agar tidak ada siswa stres yang terlewat, serta memastikan bahwa intervensi yang diberikan tepat sasaran tanpa terlalu banyak siswa tidak stres yang salah teridentifikasi.



Gambar 4. *Confusion matrix* model terbaik dari *Random Forest*

Pada Gambar 4 di atas, model *Random Forest* dengan metode BoW menghasilkan performa yang baik dalam memprediksi stres siswa. Dari total 369 data, model ini berhasil membuat 345 prediksi benar dengan 153 siswa yang tidak stres diprediksi dengan benar (*True Negative*) dan 192 siswa yang benar-benar stres juga teridentifikasi dengan tepat (*True Positive*). Namun, terdapat beberapa kesalahan prediksi yaitu 13 kasus *False Positive*, di mana siswa yang sebenarnya tidak stres diprediksi sebagai stres dan 11 kasus *False Negative*, di mana siswa yang mengalami stres tidak terdeteksi oleh model. Kesalahan *False Negative* ini cukup penting untuk diperhatikan karena berisiko melewatkan siswa yang membutuhkan bantuan. Meskipun demikian, secara keseluruhan model ini menunjukkan performa yang kuat dengan kesalahan prediksi yang relatif kecil, tetapi tetap memerlukan peningkatan untuk mengurangi kasus *False Negative* dan *False Positive*.



Gambar 5. *Confusion matrix* model terbaik dari *SVM*

Dan pada Gambar 5 di atas, model *SVM* dengan metode BoW menunjukkan performa yang cukup baik dalam memprediksi stres siswa. Dari total 369 data, model ini berhasil membuat 343 prediksi yang benar, dengan 152 siswa yang tidak stres diprediksi dengan benar (*True Negative*) dan 191 siswa yang stres teridentifikasi dengan tepat (*True Positive*). Hal ini menunjukkan bahwa model mampu mengenali pola stres dengan cukup akurat. Namun, terdapat 14 kasus *False Positive*, di mana siswa yang tidak stres diprediksi sebagai stres. Meskipun dampaknya tidak terlalu besar, kesalahan ini dapat menyebabkan intervensi yang tidak diperlukan. Di sisi lain, ada 12 kasus *False Negative*, di mana siswa yang mengalami stres tidak terdeteksi, yang lebih berisiko karena siswa yang membutuhkan perhatian terabaikan. Meskipun hasil model *SVM* dengan BoW ini cukup baik, masih ada ruang untuk perbaikan, terutama dalam mengurangi jumlah kesalahan prediksi. Evaluasi lebih lanjut dan eksperimen dengan teknik lain diharapkan dapat meningkatkan akurasi model untuk mendeteksi stres siswa dengan lebih tepat..

Tabel 4. *Classification report* model terbaik *Logistic Regression* pada data uji

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
0	0,93	0,92	0,92	166
1	0,93	0,94	0,94	203
<i>Accuracy</i>			0,93	369
<i>Macro Avg</i>	0,93	0,93	0,93	369
<i>Weighted Avg</i>	0,93	0,93	0,93	369

Berdasarkan *Classification report* yang terlihat pada Tabel 4 di atas, model *Logistic Regression* dengan metode BoW menunjukkan *accuracy* sebesar 93% dalam data uji. Dengan *precision* di angka 0,93 serta *recall* dan *F1-score* antara 0,92 sampai 0,94 untuk kedua kelas (stres dan tidak stres). Hal ini memperlihatkan bahwa *Logistic Regression* mampu memprediksi stres secara konsisten, meskipun perlu peningkatan dalam mengidentifikasi kasus *False Negative* untuk mengurangi risiko siswa stres yang terlewatkan. Hasil ini juga berbanding terbalik dengan kegunaan *Logistic Regression* yang seharusnya bisa lebih baik dalam menangani data biner.

Tabel 5. *Classification report* model terbaik *Naive Bayes* pada data uji

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
0	0,95	0,96	0,95	162
1	0,97	0,96	0,96	207
<i>Accuracy</i>			0,96	369
<i>Macro Avg</i>	0,96	0,96	0,96	369
<i>Weighted Avg</i>	0,96	0,96	0,96	369

Pada Tabel 5 di atas, *Naive Bayes* dengan fitur *Bigram* menghasilkan performa terbaik, dengan *accuracy* mencapai 96% pada data uji. Model ini memiliki *precision* antara 0,95 sampai 0,97, serta *recall* yang tinggi pada kedua kelas, menunjukkan bahwa *Naive Bayes* sangat efektif dalam menangkap pola teks yang kompleks dan relevan. *F1-score* yang mencapai 0,96 menunjukkan bahwa model ini sangat akurat dan andal dalam memprediksi stres pada siswa. Tingginya *precision* dan *recall* menandakan bahwa model ini berhasil meminimalkan jumlah kesalahan prediksi, terutama pada kasus siswa yang benar-benar stres.

Tabel 6. *Classification report* model terbaik *Random Forest* pada data uji

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
0	0,93	0,92	0,93	166
1	0,94	0,95	0,94	203
<i>Accuracy</i>			0,93	369
<i>Macro Avg</i>	0,93	0,93	0,93	369
<i>Weighted Avg</i>	0,93	0,93	0,93	369

Pada Tabel 6 di atas, model *Random Forest* dengan metode BoW memiliki *accuracy* sebesar 93%, serta *precision*, *recall*, dan *F1-score* yang berada di sekitar 0,92 sampai 0,95 untuk kedua kelas. Model ini mampu mendeteksi stres dengan tingkat *recall* yang sedikit lebih tinggi dibandingkan model lain, sehingga memberikan keunggulan dalam menemukan siswa yang benar-benar stres.

Tabel 7. *Classification report* model terbaik *SVM* pada data uji

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
0	0,93	0,92	0,92	166
1	0,93	0,94	0,94	203
<i>Accuracy</i>			0,93	369
<i>Macro Avg</i>	0,93	0,93	0,93	369
<i>Weighted Avg</i>	0,93	0,93	0,93	369

Dan pada Tabel 7 di atas, menunjukkan bahwa model *SVM* dengan metode BoW mencapai *accuracy* 93%, dengan *precision*, *recall*, dan *F1-score* model ini berada di kisaran 0,92 sampai 0,94, menunjukkan kemampuan yang cukup baik dalam prediksi secara keseluruhan. Meskipun performanya mirip dengan *Logistic Regression*, *SVM* memberikan hasil yang lebih konsisten pada kedua kelas, sehingga cocok untuk aplikasi yang memerlukan keseimbangan antara *precision* dan *recall*.

Dalam dataset pengujian, evaluasi *classification report* untuk model *Logistic Regression*, *Naive Bayes*, *Random Forest* dan *SVM* menunjukkan perbedaan kinerja yang signifikan di antara model-model tersebut. Model

Naive Bayes dengan *Bigram* mencapai kinerja tertinggi dengan akurasi 96%, *precision* 0,95 – 0,97, dan *recall* yang tinggi untuk kedua kelas, menunjukkan bahwa *Naive Bayes* efektif dalam menangkap pola teks meskipun menggunakan teknik probabilistik yang sederhana. Sebagai perbandingan, *Logistic Regression* dan SVM, meskipun dirancang baik untuk data biner, keduanya menunjukkan akurasi yang sedikit lebih rendah, yaitu 93% dengan ekstraksi fitur BoW, serta *precision*, *recall*, dan *F1-score* yang serupa di antara 0,92 hingga 0,94. *Random Forest*, meskipun dikenal untuk meminimalkan risiko *overfitting*, justru menunjukkan performa yang sebanding dengan SVM dan *Logistic Regression*, dengan akurasi 93% dengan ekstraksi fitur yang sama yaitu BoW. Meskipun memiliki *precision* dan *recall* yang sedikit lebih baik untuk kelas positif (1), *Random Forest* menunjukkan hasil yang sangat mirip dengan kedua model tersebut dalam metrik evaluasi lainnya, seperti *F1-score*. Dalam beberapa kasus, *Random Forest* juga menunjukkan tanda-tanda *overfitting* yang menyebabkan sedikit penurunan konsistensi prediksi dibandingkan dengan model lainnya. Singkatnya, meskipun model-model seperti *Logistic Regression*, SVM dan *Random Forest* menunjukkan performa yang baik, *Naive Bayes* dengan *Bigram* tetap terbukti sebagai model yang lebih unggul dalam menganalisis teks pada dataset ini, dengan kemampuan untuk menghasilkan prediksi yang lebih akurat dan konsisten.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi stres siswa SMA dengan memanfaatkan teknik NLP dengan berbagai ekstraksi fitur dan algoritma *Machine Learning*. Berdasarkan hasil eksperimen yang dilakukan, beberapa model seperti *Naive Bayes*, *Logistic Regression*, *Random Forest*, dan SVM menunjukkan performa yang baik dalam memprediksi kondisi stres dengan tingkat akurasi yang bervariasi. Pemilihan kombinasi antara ekstraksi fitur dan algoritma memiliki dampak signifikan terhadap akurasi dan keandalan model. Model terbaik dalam penelitian ini adalah *Naive Bayes* dengan fitur *Bigram*, yang menghasilkan akurasi sebesar 95,6%. Hal ini menunjukkan bahwa pendekatan ini sangat efektif dalam mengidentifikasi stres pada siswa, dengan kemampuan untuk mendeteksi pola-pola teks yang kompleks yang berhubungan dengan stres. Model ini menonjol dalam hal *precision* dan *recall*, yang menunjukkan kemampuannya dalam meminimalkan kesalahan prediksi, baik *False Positive* maupun *False Negative*. Namun, meskipun sebagian besar model menunjukkan hasil yang sangat baik, masih terdapat tantangan yang perlu diatasi, seperti mengurangi kesalahan prediksi, terutama *False Negative* yang dapat berisiko melewatkan siswa yang membutuhkan intervensi. Oleh karena itu, untuk meningkatkan akurasi model, penting untuk mengoptimalkan parameter dan algoritma, serta melakukan validasi yang lebih mendalam dengan data yang lebih bervariasi. Secara keseluruhan, penelitian ini memberikan bukti bahwa metode NLP dan *Machine Learning* memiliki potensi yang sangat besar untuk diterapkan dalam mendeteksi stres di kalangan siswa, dengan hasil yang memadai untuk digunakan dalam aplikasi nyata. Namun, penelitian lebih lanjut diperlukan untuk meningkatkan kemampuan generalisasi model dan mengurangi kesalahan prediksi agar dapat memberikan intervensi yang lebih tepat sasaran dalam mendukung kesehatan mental siswa.

REFERENCES

- [1] A. Rahma and S. Cahyani, "Analisa Hubungan Antara Rasa Syukur Terhadap Kesehatan Mental Remaja di Banjarbaru," *Jurnal Religion: Jurnal Agama, Sosial, dan Budaya*, vol. 1, no. 6, 2023, [Online]. Available: <https://maryamsejahtera.com/index.php/Religion/index>
- [2] N. Fulambarkar, B. Seo, A. Testerman, M. Rees, K. Bausback, and E. Bunge, "Review: Meta-analysis on mindfulness-based interventions for adolescents' stress, depression, and anxiety in school settings: a cautionary tale," May 01, 2023, *John Wiley and Sons Inc.* doi: 10.1111/camh.12572.
- [3] A. Z. B. Mentari, E. Liana, and T. Y. R. Pristya, "Teknik Manajemen Stres yang Paling Efektif pada Remaja: Literature Review," *JURNAL ILMIAH KESEHATAN MASYARAKAT: Media Komunikasi Komunitas Kesehatan Masyarakat*, vol. 12, no. 4, pp. 191–196, Dec. 2020, doi: 10.52022/jikm.v12i4.69.
- [4] R. Syahputra and N. P. Siregar, "HUBUNGAN ANTARA TINGKAT STRES DENGAN KEJADIAN SINDROM DISPEPSIA FUNGSIONAL PADA MAHASISWA FAKULTAS KEDOKTERAN UNIVERSITAS ISLAM SUMATERA UTARA TAHUN 2020," *Jurnal Kedokteran Ibnu Nafis*, vol. 10, no. 2, pp. 101–109, Dec. 2021, doi: 10.30743/jkin.v10i2.178.
- [5] M. A. Setiawan, R. W. Eriyanti, A. M. Huda, and A. Rofieq, "Problematisasi Tuntutan Eksternal bagi Kesejahteraan Psikologis Peserta Didik: Literature Review," *JURNAL BIMBINGAN DAN KONSELING AR-RAHMAN*, vol. 10, no. 1, p. 144, Jun. 2024, doi: 10.31602/jbkr.v10i1.13719.
- [6] F. Yoduke, N. H. C. Daulima, and M. Mustikasari, "STRATEGI KOPING PADA REMAJA DALAM MENURUNKAN GEJALA KECEMASAN DAN DEPRESI: Literature Review," *Alauddin Scientific Journal of Nursing*, vol. 4, no. 1, pp. 16–24, Apr. 2023, doi: 10.24252/asjn.v4i1.34251.
- [7] M. Malgaroli, T. D. Hull, J. M. Zech, and T. Althoff, "Natural language processing for mental health interventions: a systematic review and research framework," *Transl Psychiatry*, vol. 13, no. 1, p. 309, Oct. 2023, doi: 10.1038/s41398-023-02592-2.
- [8] M. Amien, "Sejarah dan Perkembangan Teknik Natural Language Processing (NLP) Bahasa Indonesia: Tinjauan tentang sejarah, perkembangan teknologi, dan aplikasi NLP dalam bahasa Indonesia," Mar. 2023, doi: 10.48550/arXiv.2304.02746.
- [9] E. Gutierrez, W. Karwowski, K. Fiok, M. R. Davahli, T. Liciaga, and T. Ahram, "Analysis of Human Behavior by Mining Textual Data: Current Research Topics and Analytical Techniques," *Symmetry (Basel)*, vol. 13, no. 7, p. 1276, Jul. 2021, doi: 10.3390/sym13071276.
- [10] C. Su, Z. Xu, J. Pathak, and F. Wang, "Deep learning in mental health outcome research: a scoping review," *Transl Psychiatry*, vol. 10, no. 1, p. 116, Apr. 2020, doi: 10.1038/s41398-020-0780-3.



- [11] B. Lamichhane, "Evaluation of ChatGPT for NLP-based Mental Health Applications," Mar. 2023, doi: <https://doi.org/10.48550/arXiv.2303.15727>.
- [12] A. Le Glaz *et al.*, "Machine Learning and Natural Language Processing in Mental Health: Systematic Review," *J Med Internet Res*, vol. 23, no. 5, p. e15708, May 2021, doi: 10.2196/15708.
- [13] S. Sinha, "EARLY STRESS DETECTION USING NATURAL LANGUAGE PROCESSING AND MACHINE LEARNING," vol. 13, Jan. 2024, doi: 10.36106/ijrsr.
- [14] K. Kumari and S. Das, "Stress Detection System using Natural Language Processing and Machine Learning Techniques," 2022. [Online]. Available: <https://iiitranchi.ac.in/>
- [15] S. A. Rajagukguk, "TINJAUAN PUSTAKA SISTEMATIS: PREDIKSI PRESTASI BELAJAR PESERTA DIDIK DENGAN ALGORITMA PEMBELAJARAN MESIN," *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, vol. 1, no. 1, Aug. 2021, doi: 10.20885/snati.v1i1.4.
- [16] D. M. Low, L. Rumker, T. Talkar, J. Torous, G. Cecchi, and S. S. Ghosh, "Natural Language Processing Reveals Vulnerable Mental Health Support Groups and Heightened Health Anxiety on Reddit During COVID-19: Observational Study," *J Med Internet Res*, vol. 22, no. 10, p. e22635, Oct. 2020, doi: 10.2196/22635.
- [17] M. Najamuddin, D. Miharja, and S. Adhkar, "Implementasi Chatbot Deteksi Depresi Dini Pada Mahasiswa Dengan Phq-9 (Patient Health Questionnaire) Menggunakan NLP (Natural Language Processing)," *Pelita Bangsa*, vol. 1, no. 1, 2022, Accessed: Oct. 13, 2024. [Online]. Available: <https://jurnal.pelitabangsa.ac.id/index.php/SAINTEK/article/view/1156>
- [18] N. H. Kim, J. M. Kim, D. M. Park, S. R. Ji, and J. W. Kim, "Analysis of depression in social media texts through the Patient Health Questionnaire-9 and natural language processing," *Digit Health*, vol. 8, 2022, doi: 10.1177/20552076221114204.
- [19] P. Syahputra and R. Kurniawan, "Analisis Sentimen Terhadap Kesehatan Mental Remaja Menggunakan Metode Naive Bayes," *Journal of Information System Research (JOSH)*, vol. 5, no. 4, pp. 1216–1224, 2024, doi: 10.47065/josh.v5i4.5644.
- [20] M. Bader, M. Abdelwanis, M. Maalouf, and H. F. Jelinek, "Detecting depression severity using weighted random forest and oxidative stress biomarkers," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-67251-y.
- [21] A. Van Fadhila *et al.*, "Implementasi Metode Machine Learning Untuk Mendeteksi Tingkat Stres Manusia Berdasarkan Kualitas Tidur," 2023. Accessed: Nov. 07, 2024. [Online]. Available: <https://conference.upnvj.ac.id/index.php/senamika/article/view/2415>
- [22] A. Heins, "School Anxiety: Signs and Strategies," Anxiety and Depression Association of America (ADAA). Accessed: Oct. 24, 2024. [Online]. Available: <https://adaa.org/learn-from-us/from-the-experts/blog-posts/consumer/school-anxiety-signs-and-strategies>
- [23] M. M. A. Smith, "Burnout: Symptoms, Treatment, and Coping Strategy Tips," HelpGuide.org. Accessed: Oct. 24, 2024. [Online]. Available: <https://www.helpguide.org/mental-health/stress/burnout-prevention-and-recovery>
- [24] A. Pietrangelo, "The Effects of Stress on Your Body," Healthline. Accessed: Oct. 24, 2024. [Online]. Available: <https://www.healthline.com/health/stress/effects-on-body>
- [25] Mayo Clinic Staff, "Stress management," Mayo Clinic. Accessed: Oct. 24, 2024. [Online]. Available: <https://www.mayoclinic.org/healthy-lifestyle/stress-management/in-depth/stress/art-20046037>
- [26] E. Scott, "How to Deal With Frustration," Verywell mind. Accessed: Oct. 24, 2024. [Online]. Available: <https://www.verywellmind.com/feel-less-frustrated-when-stressed-3145200>
- [27] F. Kasmin, "Stress Detection Through Text in Social Media Using Machine Learning Techniques," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, pp. 162–176, Oct. 2024, doi: 10.37934/araset.63.2.162176.