

Analisis Komparatif Algoritma Machine Learning untuk Prediksi Depresi Mahasiswa Berbasis PHQ-9

Asep Setiyono¹, Yulia Wardani², Margaretha Kurniastuti³

¹Keperawatan, Kesehatan, Sekolah Tinggi Ilmu Kesehatan Panti Rapih Yogyakarta, Indonesia

Email: ¹*robertus.setiyono@stikespantirapih.ac.id, ²yulia_wardani@stikespantirapih.ac.id,

³margarethakurniastuti@stikespantirapih.ac.id

Email Penulis Korespondensi: robertus.setiyono@stikespantirapih.ac.id*

Submitted: 02/05/2026; Accepted: 08/06/2026; Published: 30/06/2026

Abstrak-Depresi adalah suatu kondisi gangguan mental yang perlu mendapatkan perhatian serius, terutama di kalangan mahasiswa yang menghadapi tekanan akademik dan sosial. Prevalensi depresi pada mahasiswa mencapai angka yang mengkhawatirkan, dengan deteksi dini dan intervensi yang akurat menjadi krusial. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja empat algoritma *machine learning* (*Random Forest*, *Logistic Regression*, *Support Vector Machine*, dan *Extreme Gradient Boosting*) dalam memprediksi depresi mahasiswa di Sekolah Tinggi Ilmu Kesehatan Panti Rapih Yogyakarta. Menggunakan *dataset* kuesioner PHQ-9 dari 679 mahasiswa, data dipra-proses termasuk penanganan ketidakseimbangan kelas menggunakan SMOTE pada data pelatihan. Evaluasi model dilakukan dengan *Stratified K-Fold Cross Validation* dan metrik seperti akurasi, *presisi*, *recall*, *F1-Score* serta *ROC AUC*. Hasil menunjukkan bahwa semua model memiliki kinerja prediktif yang sangat tinggi (akurasi > 92%). *Support Vector Machine* menunjukkan kinerja terbaik secara keseluruhan dengan akurasi 94,30%, *F1-Score* 94,40% dan *ROC AUC* 98,57%. *Random Forest* unggul dalam *recall* sebesar 96,12% sementara *Extreme Gradient Boosting* mencapai *presisi* tertinggi (93,51%). Penelitian ini mengidentifikasi *Support Vector Machine* sebagai algoritma paling seimbang untuk prediksi depresi mahasiswa, menawarkan potensi sebagai alat skrining awal yang objektif dan efisien.

Kata Kunci: *machine learning*; prediksi; depresi; PHQ-9; komparatif

Abstract-Depression is a serious mental health problem, especially among college students facing academic and social pressures. The prevalence of depression in college students is alarming, making early detection and accurate intervention crucial. This study aims to analyze and compare the performance of four machine learning algorithms (*Random Forest*, *Logistic Regression*, *Support Vector Machine*, and *Extreme Gradient Boosting*) in predicting depression in college students at Panti Rapih Health Sciences College, Yogyakarta. Using the PHQ-9 questionnaire dataset from 679 college students, the data was preprocessed, including class imbalance handling using SMOTE on the training data. Model evaluation was conducted using *Stratified K-Fold Cross Validation* and metrics such as accuracy, precision, recall, *F1-Score*, and *ROC AUC*. The results showed that all models had very high predictive performance (accuracy > 92%). *Support Vector Machine* showed the best overall performance with an accuracy of 94.30%, an *F1-Score* of 94.40%, and a *ROC AUC* of 98.57%. *Random Forest* excelled in recall with 96.12%, while *Extreme Gradient Boosting* achieved the highest precision (93.51%). This study identified *Support Vector Machine* as the most balanced algorithm for predicting depression in college students, offering potential as an objective and efficient initial screening tool.

Keywords: machine learning; prediction; depression; PHQ-9; comparative

1. PENDAHULUAN

Depresi merupakan gangguan kesehatan mental serius yang berdampak pada jutaan individu secara global, berkontribusi signifikan terhadap beban penyakit dunia [1]. Kondisi ini tidak hanya menimbulkan penderitaan emosional mendalam bagi individu yang mengalaminya, namun dampak yang timbul tidak hanya sebatas itu, terdapat pula implikasi sosial dan ekonomi yang meluas di antaranya adalah berkurangnya efisiensi dalam pekerjaan dan bertambahnya tekanan terhadap fasilitas atau sistem pelayanan kesehatan, serta dampak negatif terhadap kualitas hidup secara menyeluruh, sehingga menempatkannya sebagai prioritas utama dalam agenda kesehatan masyarakat global [2]. Sebagai salah satu gangguan psikologis paling sering terjadi, depresi umumnya dikenali dari suasana hati yang terus-menerus menurun, perasaan sedih yang tak kunjung hilang, serta perasaan putus asa, bersalah, dan tidak berharga yang mempengaruhi individu [2]. Manifestasi gejala-gejala inti ini dapat sangat bervariasi antar individu, namun secara kolektif, mereka cenderung mengganggu fungsi kognitif seperti konsentrasi dan pengambilan keputusan, menghambat motivasi untuk interaksi sosial, dan secara drastis menurunkan kemampuan seseorang untuk menjalankan aktivitas harian, baik dalam lingkup pribadi maupun profesional, termasuk performa akademik bagi mahasiswa. Studi-studi global menunjukkan bahwa prevalensi depresi di kalangan mahasiswa perguruan tinggi mencapai angka yang mengkhawatirkan, dengan laporan yang mengindikasikan sekitar 20% hingga 45% mahasiswa mengalami gejala depresi dalam derajat yang bervariasi selama masa studi mereka [3]. Di Indonesia, data serupa juga memperlihatkan bahwa mahasiswa sering kali menghadapi tekanan akademik, tuntutan sosial, dan ketidakpastian masa depan yang dapat memicu atau memperparah kondisi depresi, dengan prevalensi yang berkisar antara 22% hingga 33% di berbagai institusi pendidikan tinggi [4], [5]. Angka ini selaras dengan laporan terkini di Daerah Istimewa Yogyakarta (DIY) yang

menunjukkan prevalensi depresi pada populasi mahasiswa mencapai sekitar 24,6%, mengindikasikan bahwa masalah kesehatan mental ini merupakan isu krusial yang memerlukan perhatian serius dan intervensi yang komprehensif di tingkat regional [6]. Secara khusus di STIKes Panti Rapih Yogyakarta, meskipun data insiden depresi yang spesifik belum secara formal terdokumentasi secara komprehensif, observasi awal dari konselor kampus dan laporan anekdotal dari dosen serta mahasiswa menunjukkan adanya peningkatan kebutuhan akan dukungan kesehatan mental. Rentang keparahan depresi yang luas ini menggarisbawahi pentingnya deteksi dini dan intervensi yang terpersonalisasi. Saat ini, metode deteksi depresi yang dominan masih mengandalkan kuesioner *self report* dan wawancara klinis yang meskipun esensial, memiliki keterbatasan dalam hal subjektivitas, konsistensi antar rater, serta membutuhkan waktu dan sumber daya yang signifikan, sehingga seringkali belum mampu mengidentifikasi depresi pada tahap awal secara proaktif dan dalam skala besar. Mengingat kompleksitas dan variasi gejala depresi, serta urgensi penanganannya yang lebih serius, pengembangan metode deteksi yang objektif, akurat, dan efisien menjadi sangat krusial, terutama di lingkungan padat seperti institusi pendidikan tinggi.

Menyikapi kompleksitas dan urgensi penanganan depresi, studi ini menyarankan implementasi teknologi *Machine Learning* (ML) sebagai solusi yang objektif, akurat, dan efisien. ML adalah ilmu yang mempelajari algoritma komputer yang bisa mengenali pola-pola di dalam data dengan tujuan untuk mengubah macam data menjadi satu tindakan nyata dengan sedikit campur tangan dari manusia [7]. *Artificial Intelligence* (AI) menghadirkan solusi yang sangat menjanjikan dengan cara mengotomatisasi dan meningkatkan ketepatan proses diagnostik depresi melalui penerapan pembelajaran mesin. Sistem ini memiliki kemampuan untuk menganalisis volume data yang besar guna mengidentifikasi pola-pola spesifik yang berkaitan dengan kondisi depresi [8]. Algoritma ML yang mampu mengembangkan model untuk keperluan prediksi sudah menunjukkan efektivitas yang tinggi dalam mengatasi beragam isu atau masalah terkait klasifikasi [9]. Adopsi teknologi pembelajaran mesin dalam skrining depresi memungkinkan analisis data yang lebih mendalam dan identifikasi indikator yang mungkin terlewatkan oleh metode diagnostik konvensional, sehingga berpotensi mempercepat proses identifikasi individu yang berisiko, memungkinkan intervensi lebih dini, dan pada akhirnya, berkontribusi pada peningkatan kesehatan mental mahasiswa secara signifikan. Penelitian ini secara khusus berfokus pada mahasiswa di STIKes Panti Rapih Yogyakarta menggunakan *dataset* depresi dan melakukan analisis kinerja komparatif terhadap empat model pembelajaran mesin yang berbeda: *Random Forest*, *Logistic Regression*, *Support Vector Machine*, dan *Extreme Gradient Boosting*. *Dataset* yang berasal dari pengalaman yang telah dikumpulkan, diolah dan diubah menjadi bentuk yang dapat diproses oleh mesin [10]. Melalui komparasi ini, penelitian akan mengidentifikasi model mana yang paling optimal dan valid untuk memprediksi depresi secara akurat di kalangan mahasiswa STIKes Panti Rapih Yogyakarta, menawarkan keunggulan melalui pendekatan yang lebih terstruktur dan spesifik.

Berbagai penelitian terkini telah menunjukkan upaya untuk meningkatkan deteksi depresi melalui teknologi, termasuk penggunaan teknik data mining dan *machine learning* pada berbagai jenis data. Misalnya, studi oleh [11] menggunakan data media sosial untuk memprediksi depresi, sementara [12] mengeksplorasi data sensorik dari *wearable devices*. Dalam lima tahun terakhir, beberapa penelitian terkait yang sejenis dengan penelitian ini telah mengkaji penggunaan algoritma *machine learning* untuk prediksi depresi. Penelitian dengan melakukan prediksi gangguan depresi menggunakan pendekatan *machine learning* dari data *NHANES*. Studi ini membandingkan *Logistic Regression*, *Random Forest*, *Naive Bayes*, *SVM*, *XGBoost*, dan *LightGBM*, dengan *XGBoost* mencapai akurasi tertinggi sebesar 69% [13]. Penelitian lain meneliti algoritma *machine learning* untuk memprediksi depresi pada orang dewasa yang lebih tua di Tiongkok. Penelitian ini membandingkan *Logistic Regression*, *KNN*, *SVM*, *Decision Tree*, *LightGBM*, dan *Random Forest*, dengan *LightGBM* menunjukkan akurasi tertinggi sebesar 69% [14]. Penelitian lainnya mengeksplorasi *Explainable Machine Learning* dalam prediksi depresi, menggunakan *Logistic Regression*, *SVM*, *XGBoost*, dan *Neural Network*. Hasil penelitian mereka menunjukkan bahwa *XGBoost* mencapai akurasi yang sangat tinggi, yaitu 97,8% [15]. Penelitian lain juga meneliti yang berfokus pada deteksi dan interpretasi depresi yang efektif dengan mengintegrasikan *machine learning*, *deep learning*, model bahasa, dan AI yang dapat dijelaskan *explainable AI*. Dalam penelitian ini, *Random Forest* menunjukkan akurasi tertinggi sebesar 91,1% [8]. Penelitian yang lainnya melakukan deteksi depresi anak menggunakan metode *machine learning*, membandingkan *Random Forest*, *XGBoost*, *Decision Tree*, dan *GaussianNB*. Hasil mereka menunjukkan bahwa *Random Forest* mencapai akurasi tertinggi sebesar 95% [16].

Meskipun penelitian-penelitian sebelumnya telah memperlihatkan bahwa berbagai algoritma pembelajaran mesin cukup ampuh atau berhasil dalam melakukan prediksi depresi di berbagai populasi dan menggunakan jenis data yang beragam, terdapat gap analisis yang signifikan. Secara spesifik penelitian yang ada belum secara komprehensif merepresentasikan kompleksitas depresi pada mahasiswa bidang kesehatan, terutama dalam konteks institusi pendidikan tinggi di Indonesia seperti STIKes Panti Rapih Yogyakarta. Studi-studi sebelumnya, meskipun memberikan dasar metodologis yang kuat, belum menyajikan model *machine learning* yang dioptimalkan dan disesuaikan dengan karakteristik unik populasi mahasiswa kesehatan yang dihadapkan pada tekanan akademis, tuntutan sosial dan kekhawatiran karier yang spesifik. Selain itu, perbandingan kinerja komprehensif antara model-model *machine learning* utama (*Random Forest*, *Logistic Regression*, *Support Vector Machine*, dan *Extreme Gradient Boosting*) yang disesuaikan dengan *dataset* lokal dan karakteristik demografi mahasiswa STIKes Panti Rapih Yogyakarta masih terbatas. Maka dari itu studi ini dilaksanakan untuk mengisi atau mengatasi kesenjangan yang ada, melalui pendekatan yang sistematis menganalisis kinerja komparatif keempat algoritma tersebut, untuk

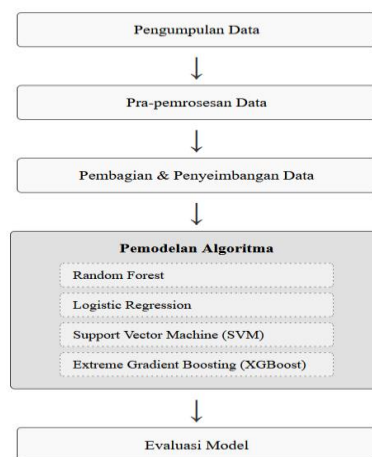
mengidentifikasi model yang paling optimal dan sesuai dalam konteks populasi yang spesifik ini, sehingga hasil yang diperoleh dapat memberikan rekomendasi yang lebih kuat dan dapat diimplementasikan.

Berdasarkan latar belakang dan analisis gap tersebut, penelitian ini memiliki tujuan utama sebagai berikut yaitu mengetahui model algoritma ML yang spesifik untuk memprediksi depresi pada mahasiswa STIKes Panti Rapih Yogyakarta dan juga mengidentifikasi algoritma ML yang menunjukkan tingkat akurasi prediksi tertinggi dalam memprediksi depresi pada mahasiswa STIKes Panti Rapih Yogyakarta. Penelitian ini diharapkan dapat memberikan kontribusi signifikan berupa pembentukan model ML yang spesifik untuk prediksi depresi mahasiswa dan mengidentifikasi algoritma prediktif terbaik. Selain itu, hasil penelitian ini diharapkan dapat menyediakan landasan yang kokoh sebagai fondasi untuk penciptaan instrumen atau alat pendeteksi depresi yang lebih akurat dan efisien, mendukung deteksi dini, dan pada akhirnya, meningkatkan upaya penanganan kesehatan mental mahasiswa di STIKes Panti Rapih Yogyakarta secara lebih proaktif dan terpersonalisasi.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Prosedur penelitian ini mengikuti alur sistematis yang dapat digambarkan sebagai serangkaian tahapan berurutan, dimulai dari persiapan data hingga evaluasi model, seperti pada *flowchart* Gambar 1.



Gambar 1. Tahapan Penelitian

2.2 Pengumpulan Data

Tahap ini melibatkan pengumpulan data primer kuesioner *PHQ-9* dari 679 mahasiswa STIKes Panti Rapih Yogyakarta untuk tahun akademik 2025/2026. Penelitian ini mengimplementasikan teknik *total sampling*, yaitu ketika setiap anggota dari populasi dijadikan objek atau partisipan dalam studi untuk mendapatkan data yang komprehensif. Instrumen utama yang digunakan adalah kuesioner *PHQ-9* yang telah diadaptasi, serta mencakup pengumpulan data demografi. *PHQ-9* merupakan instrumen skrining depresi yang divalidasi terdiri dari 9 pertanyaan dengan skala *likert* 0-3, menghasilkan skor total 0-27. *Cut off* skor 10 atau lebih digunakan untuk mengindikasikan depresi. Sebelum pengumpulan data, peneliti telah memperoleh persetujuan etik dan *informed consent* dari setiap responden, memastikan bahwa hak-hak partisipan dihormati dan kerahasiaan data terjaga.

2.3 Pra-Pemrosesan Data

Langkah ini krusial untuk mengolah data awal supaya memenuhi persyaratan algoritma *machine learning*. Proses ini bertujuan untuk menjamin bahwa data menjadi rapi tertata dengan baik dan telah siap untuk digunakan dalam proses pelatihan model. Ini mencakup penghapusan kolom yang tidak relevan, seperti kolom 'No' yang berfungsi sebagai pengidentifikasi unik dan tidak memiliki nilai prediktif. Selain itu, teknik *One-Hot Encoding* diterapkan pada fitur *kategorikal* seperti Jenis_Kelamin dan Program_Studi untuk mengubahnya menjadi format biner. Hal ini dilakukan untuk mencegah algoritma salah menginterpretasikan urutan nilai dan menghindari masalah multikolinearitas, memastikan setiap kategori memiliki representasi independen.

2.4 Pembagian dan Penyeimbangan Data

Sesudah data selesai melewati fase pra-pemrosesan, kumpulan data tersebut selanjutnya akan dipisahkan menjadi dua segmen: 80% akan dipakai sebagai data untuk melatih model (data pelatihan), sementara 20% sisanya akan digunakan untuk menguji model (data pengujian). Pembagian ini dilakukan secara acak dengan mempertahankan

rasio kelas target pada setiap *subset* untuk menjaga representasi proporsional. Tahap ini juga menangani masalah ketidakseimbangan kelas pada variabel target 'status_depresi'. Untuk mengatasi bias yang mungkin muncul akibat proporsi kelas yang tidak seimbang (misalnya, lebih banyak kelas 'tidak depresi' daripada 'depresi'), teknik *Synthetic Minority Over-sampling Technique (SMOTE)* diterapkan secara khusus pada data pelatihan. SMOTE berfungsi dengan cara menciptakan contoh-contoh data buatan untuk kategori minoritas yang bertujuan untuk menyamakan atau menyeimbangkan kuantitas sampel antara kategori mayoritas dan minoritas, memastikan model mendapatkan representasi yang cukup dari kedua kelas saat proses pelatihan.

2.5 Pemodelan Algoritma

Pada tahap ini, empat algoritma *machine learning* yang berbeda diimplementasikan dan dilatih menggunakan data pelatihan yang telah diskalakan dan diseimbangkan. Setiap algoritma memiliki karakteristik dan cara kerja tersendiri dalam mempelajari pola dari data PHQ-9 dan demografi untuk memprediksi status depresi mahasiswa. Algoritma yang digunakan meliputi:

Random Forest adalah model pembelajaran mesin *ensemble* yang efektif untuk klasifikasi dan prediksi. Algoritma ini membangun banyak pohon keputusan (*decision trees*) secara independen, masing-masing dilatih dengan subset data dan fitur acak, kemudian menggabungkan hasil prediksinya melalui *voting* atau rata-rata untuk keputusan akhir. Keunggulannya meliputi kemampuan mengatasi data tidak seimbang, nilai hilang, dan mengurangi *overfitting* [17]. Pada penelitian ini, *Random Forest* diimplementasikan dengan parameter *default sklearn*, termasuk $n_estimators=100$ (jumlah pohon keputusan) dan *criterion='gini'* (fungsi untuk mengukur kualitas pembelahan). Algoritma ini juga telah terbukti efektif untuk prediksi depresi mahasiswa dengan akurasi tinggi [18].

Logistic Regression merupakan pendekatan statistik yang digunakan untuk meninjau korelasi antara variabel-variabel bebas dengan variabel terikat yang bersifat biner atau memiliki dua kelompok. Metode ini berfungsi dengan cara mentransformasi gabungan linear dari prediktor-prediktor menjadi sebuah nilai probabilitas yang berada dalam rentang 0 hingga 1 menggunakan fungsi logit atau sigmoid, sehingga dapat memodelkan kemungkinan terjadinya suatu peristiwa [19]. Umumnya digunakan dalam tugas klasifikasi untuk memprediksi probabilitas suatu observasi termasuk dalam salah satu kelompok [20] dan efektif untuk prediksi depresi mahasiswa dengan akurasi tinggi [21]. Dalam studi ini, *Logistic Regression* diterapkan dengan parameter *default*, yaitu *solver='liblinear'* dan $C=1.0$ (*inverse of regularization strength*), yang merupakan pilihan yang baik untuk dataset kecil hingga menengah.

Support Vector Machine (SVM) ini merupakan algoritma pembelajaran mesin yang berfungsi untuk tujuan klasifikasi dan regresi. Konsep intinya adalah mencari *hyperplane* yang paling ideal, yang mampu memaksimalkan jarak (*margin*) antara dua kategori data dalam suatu ruang fitur dengan dimensi yang tinggi. SVM mengidentifikasi *support vectors* titik data terdekat dengan *hyperplane* yang krusial dalam menentukan posisi *hyperplane*. Algoritma ini fleksibel untuk masalah linier dan non-linier menggunakan teknik kernel, dikenal memiliki kemampuan generalisasi yang baik, dan efisien dalam mengelola data berdimensi tinggi [22][23]. Untuk implementasi SVM, digunakan kernel radial basis function (*RBF*) (*kernel=rbf*) dengan parameter *default* $C=1.0$ dan $\gamma=scale$. SVM juga efektif digunakan untuk prediksi depresi mahasiswa [24].

Extreme Gradient Boosting (XGBoost) sebuah algoritma pembelajaran mesin yang didasarkan pada metode *ensemble*, dengan memanfaatkan teknik *gradient boosting* pada struktur pohon keputusan (*decision tree*) guna mencapai ketepatan prediksi yang signifikan. Dirancang dengan fokus pada efisiensi, skalabilitas, serta portabilitas, *XGBoost* membangun model secara bertahap; dalam proses ini setiap model yang baru berfungsi untuk mengoreksi kekeliruan yang dibuat oleh model sebelumnya, sehingga mengintegrasikan banyak pembelajar lemah (*weak learners*) menjadi satu pembelajar kuat (*strong learner*). Keunggulannya mencakup komputasi paralel, penanganan nilai hilang, regularisasi untuk mencegah *overfitting*, serta optimasi memori dan kecepatan pemrosesan [25]. Dalam penelitian ini, *XGBoost* dikonfigurasi dengan $learning_rate=0.1$, $n_estimators=100$, dan $max_depth=3$, yang merupakan parameter umum untuk menyeimbangkan kecepatan dan performa. Algoritma ini juga unggul dalam memprediksi depresi mahasiswa [26].

2.6 Evaluasi Model

Tahap terakhir adalah evaluasi komprehensif terhadap kinerja masing-masing model *machine learning* yang telah dilatih. Evaluasi ini dilakukan menggunakan *Stratified K-Fold Cross Validation* dengan 5 *fold* untuk mendapatkan estimasi kinerja model yang lebih *robust* dan dapat digeneralisasi, menjaga proporsi kelas pada setiap *fold* [27], [28]. Metrik evaluasi yang digunakan meliputi akurasi, presisi, *recall*, *F1-score*, dan *ROC AUC score*. metrik-metrik ini umum dipakai dalam studi komparatif model klasifikasi untuk menilai kinerja keseluruhan serta *trade-off* antara *false positives* dan *false negatives* [29]. Selain itu *confusion matrix* dari setiap model pada *test set* ditampilkan untuk memberikan gambaran detail mengenai jenis kesalahan klasifikasi, membantu mengidentifikasi model terbaik dalam memprediksi depresi pada mahasiswa. Pemahaman yang mendalam tentang metrik evaluasi memerlukan definisi yang jelas mengenai komponen-komponen utama dalam *confusion matrix* yaitu:

- a. *True Positif (TP)*: Jumlah kasus positif yang diprediksi dengan benar.

- b. *True Negative* (TN): Jumlah kasus negatif yang diprediksi dengan benar
- c. *False Positive* (FP): Jumlah kasus negatif yang salah diprediksi sebagai positif
- d. *False Negative* (FN): Jumlah kasus positif yang salah diprediksi sebagai negatif
- e. *Accuracy* mengukur proporsi prediksi yang benar dari total prediksi.

Metrik ini memberikan gambaran umum mengenai kinerja model secara keseluruhan dalam mengklasifikasikan baik kasus depresi maupun non depresi.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Precision mengukur proporsi prediksi positif yang benar dari semua hasil yang diprediksi positif. Tingginya nilai *precision* sangat penting dalam konteks diagnosis karena meminimalkan kemungkinan kesalahan positif (salah mendiagnosis seseorang sebagai depresi)

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Recall mengukur proporsi *True Positive* dari semua kasus positif aktual. Nilai *recall* yang tinggi krusial dalam skrining depresi, memastikan bahwa sebagian besar kasus depresi yang sebenarnya dapat terdeteksi oleh model (menghindari *false negative*)

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

F1-Score menghitung rata-rata harmonik dari *precision* dan *recall*. Metrik ini sangat relevan ketika terdapat ketidakseimbangan kelas, memberikan keseimbangan penilaian antara kemampuan model untuk tidak salah diagnosis (*precision*) dan tidak melewatkan kasus (*recall*).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

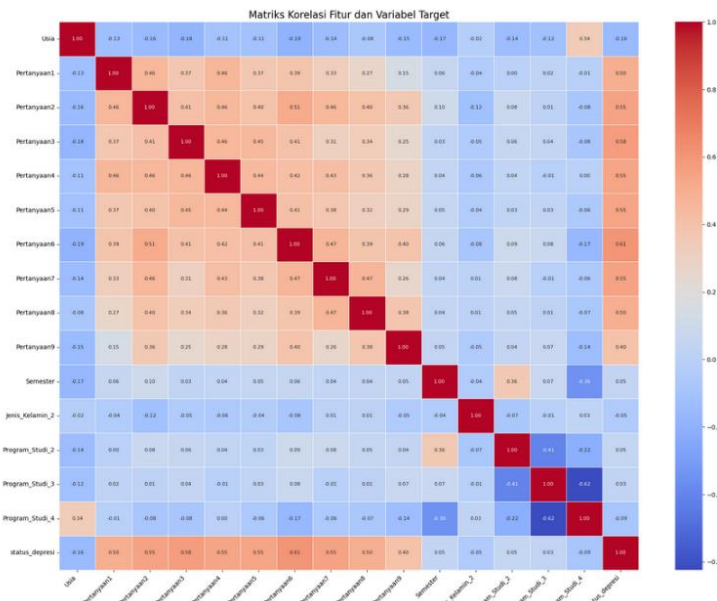
ROC AUC Mengukur kemampuan model untuk membedakan antara kelas positif dan negatif pada berbagai ambang batas [27]. Nilai *AUC* yang lebih tinggi menunjukkan kinerja model yang lebih baik. Nilai *AUC* yang mendekati 1 mengindikasikan bahwa model memiliki kemampuan diskriminasi yang sangat baik dalam membedakan subjek depresi dan non depresi, terlepas dari ambang batas klasifikasi yang dipilih.

3. HASIL DAN PEMBAHASAN

3.1. Hasil

3.1.1 Analisis Korelasi Fitur

Analisis korelasi fitur digunakan untuk memahami hubungan antara fitur-fitur dalam *dataset* dan bagaimana fitur-fitur tersebut berkorelasi dengan variabel target, dilakukan analisis korelasi menggunakan *heatmap*. Analisis korelasi ini membantu mengidentifikasi fitur mana yang memiliki hubungan linier kuat dengan depresi dan juga mendeteksi potensi multikolinearitas antar fitur. Hasil korelasi disajikan pada gambar 2.



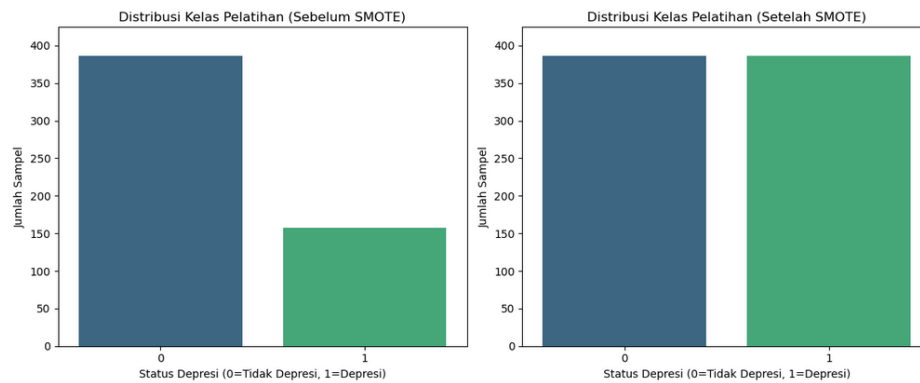
Gambar 2. Matriks Korelasi dan Variabel Target

Dari gambar 2 ini dapat dilihat bahwa fitur-fitur pertanyaan *PHQ-9* (pertanyaan 1 sampai 9) menunjukkan korelasi positif yang signifikan satu sama lain, mengindikasikan bahwa item-item ini mengukur konstruksi yang serupa yaitu gejala depresi. Beberapa pertanyaan *PHQ-9* juga menunjukkan korelasi positif yang cukup kuat dengan

variabel target ‘status_depresi’, mengindikasikan bahwa respons terhadap pertanyaan-pertanyaan ini adalah prediktor penting untuk status depresi. Sebaliknya fitur demografi seperti usia, semester, jenis-kelamin dan program_studi menunjukkan korelasi yang relatif rendah dengan ‘status_depresi’. Hal ini menunjukkan bahwa faktor demografi dalam *dataset* ini tidak memiliki hubungan linear yang kuat dengan status depresi dibandingkan dengan gejala *PHQ-9* itu sendiri. Meskipun korelasi linear rendah, tidak berarti fitur demografi tidak relevan sama sekali, namun peran mereka mungkin lebih kompleks atau non-linear.

3.1.2 Penanganan Ketidakseimbangan Kelas (*SMOTE*)

Variabel target ‘status_depresi’ menunjukkan ketidakseimbangan kelas dimana kelas tidak depresi (0) memiliki 483 sampel dan kelas Depresi (1) memiliki 196 sampel. Ketidakseimbangan ini dengan proporsi 71% untuk kelas 0 dan 29% untuk kelas 1 dapat menyebabkan model *machine learning* cenderung bias memprediksi kelas mayoritas. Untuk mengatasi masalah ini digunakan teknik *Synthetic Minority Over Sampling Technique (SMOTE)* yang diterapkan hanya pada data pelatihan. *SMOTE* bekerja dengan menghasilkan sampel-sampel sintesis untuk kelas minoritas (depresi) sehingga jumlah sampel di kelas minoritas menjadi seimbang dengan kelas mayoritas. Pada gambar 3 berikut merupakan gambar distribusi kelas sebelum dan sesudah *SMOTE*. Sebelum *SMOTE* diterapkan pada data penelitian, distribusi kelas adalah 0=386, 1=157, setelah penerapan *SMOTE* jumlah sampel untuk kedua kelas menjadi seimbang yaitu 0=386, 1=386. Hal ini memastikan model mendapatkan representasi yang cukup dari kedua kelas saat proses pelatihan sehingga dapat belajar pola-pola yang membedakan kedua kelas secara lebih efektif dan mengurangi bias prediksi terhadap kelas mayoritas.



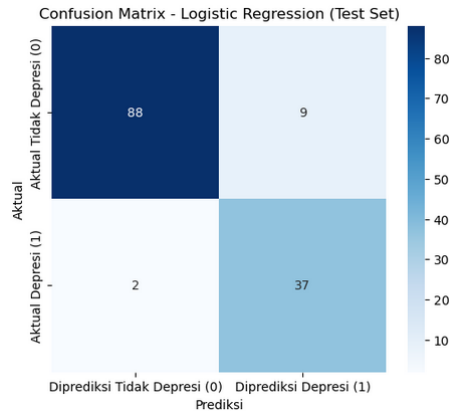
Gambar 3. Distribusi Kelas sebelum dan setelah *SMOTE*

3.1.3 Evaluasi Kinerja Model Machine Learning

Pada sub bab ini disajikan hasil evaluasi kinerja dari empat model *machine learning* yang digunakan yaitu *logistic regression*, *Random Forest*, *Support Vector Machine* dan *XGBoost*. Evaluasi utama dilakukan menggunakan *Stratified K-Fold Cross Validation* dengan 5 *fold* untuk mendapatkan estimasi kinerja model yang lebih *robust* dan dapat digeneralisasi. Metrik yang digunakan meliputi *accuracy*, *precision*, *Recall*, *F1-Score*, ROC dan AUC. Selain itu *confusion matrix* dari setiap model pada test set juga ditampilkan untuk memberikan gambaran detail mengenai jenis kesalahan klasifikasi.

a. Logistic Regression

Model *Logistic Regression* dilatih menggunakan data pelatihan yang telah diskalakan dan diseimbangkan menggunakan *SMOTE*. Evaluasi kinerja model melalui *Stratified K-Fold Cross Validation* menunjukkan hasil yang kuat: Akurasi rata-rata tercatat 0.9249, *Precision* 0.9228, *Recall* 0.9275, dan *F1-Score* 0.9250. Angka-angka ini secara kolektif mengindikasikan kemampuan model yang sangat baik dalam klasifikasi depresi. *ROC AUC Score* yang mencapai 0.9855 menegaskan kemampuan diskriminatif superior model dalam mengidentifikasi perbedaan antara mahasiswa yang menunjukkan depresi dan yang tidak, menunjukkan performa prediktif yang solid dan dapat diandalkan. Meskipun *Logistic Regression* adalah model yang relatif sederhana, kinerjanya yang tinggi menunjukkan efektivitas fitur *PHQ-9* dalam memprediksi depresi secara linier.

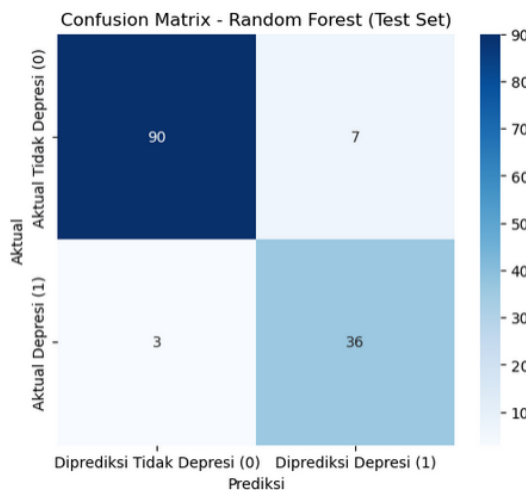


Gambar 4. Confusion Matrix Logistic Regression

Confusion matrix pada gambar 4 menunjukkan bahwa model *Logistic Regression* berhasil mengklasifikasikan 88 mahasiswa sebagai tidak depresi (*True Negatives*) dan 37 mahasiswa sebagai depresi (*True Positive*). Terdapat 9 kasus mahasiswa tidak depresi diprediksi depresi (*False Positives*) dan 2 kasus mahasiswa depresi diprediksi tidak depresi (*False Negatives*). Proporsi *False Negatives* yang rendah menunjukkan bahwa model ini cukup baik dalam menghindari kesalahan tipe II, yaitu tidak mendeteksi depresi pada individu yang sebenarnya depresi.

b. Random Forest

Selanjutnya, model *Random Forest* diimplementasikan dan dievaluasi setelah proses pelatihan pada data yang telah diskalakan dan diseimbangkan. Dari hasil *Stratified K-Fold Cross Validation*, model *Random Forest* menunjukkan kinerja prediktif yang mengesankan, Akurasi rata-rata mencapai 0.9340, menunjukkan tingkat kebenaran prediksi secara keseluruhan yang tinggi. Model ini mencatat *Precision* sebesar 0.9116, sementara keunggulannya sangat terlihat pada metrik *Recall* yang mencapai 0.9612, mengindikasikan kemampuan luar biasa dalam mengidentifikasi sebagian besar kasus depresi yang sebenarnya. Dengan *F1-Score* 0.9357, model menunjukkan keseimbangan yang kuat antara *precision* dan *recall*. Konsistensi kinerja tinggi ini diperkuat oleh *ROC AUC Score* sebesar 0.9832, menegaskan efektivitas *Random Forest* dalam membedakan antara status depresi dan non-depresi di kalangan mahasiswa. Kinerja *Random Forest* yang kuat, khususnya pada *recall*, dapat diatribusikan pada sifat *ensemble*-nya yang menggabungkan banyak *decision tree* untuk mengurangi *overfitting* dan meningkatkan generalisasi.

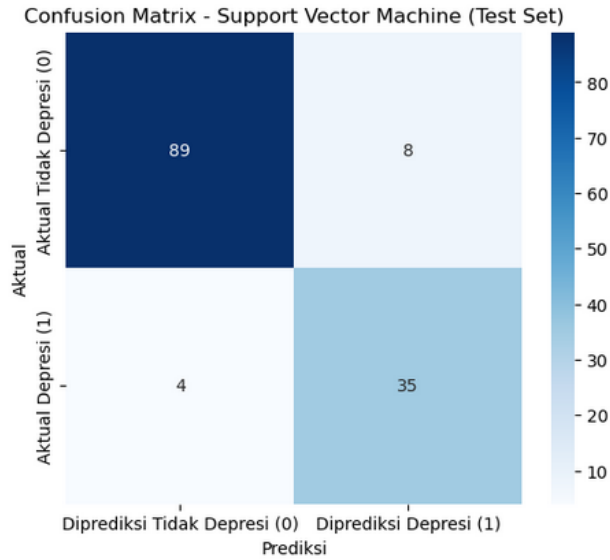


Gambar 5. Confusion Matrix Random Forest

Pada gambar 5 ini model *Random Forest* berhasil mengidentifikasi 90 mahasiswa sebagai tidak depresi (*True Negatives*) dan 36 mahasiswa sebagai depresi (*True Positive*). Model ini menghasilkan 7 kasus *False Positive* dan 3 kasus *False Negative* menunjukkan kemampuan kuat dalam mendeteksi kasus depresi (*Recall* tinggi). Nilai *Recall* yang tinggi ini sangat berharga dalam konteks skrining kesehatan mental, di mana risiko melewatkan kasus positif harus diminimalkan. Performa tinggi SVM ini konsisten dengan kemampuannya untuk menemukan *hyperplane* optimal yang memaksimalkan *margin* pemisah antar kelas, bahkan dalam ruang fitur berdimensi tinggi yang kompleks.

c. SVM

Analisis kinerja untuk model SVM juga dilakukan menggunakan *Stratified K-Fold Cross Validation* pada data pelatihan yang sudah diproses. SVM berhasil menunjukkan performa yang sangat kompetitif dengan Akurasi rata-rata tertinggi di antara semua model sebesar 0.9430. *Metric Precision*-nya mencapai 0.9301, menunjukkan kemampuan kuat dalam meminimalkan positif palsu. *Recall* model tercatat 0.9586, mengindikasikan deteksi kasus depresi yang efektif. Dengan *F1-Score* 0.944, SVM mencapai keseimbangan yang optimal antara *precision* dan *recall*. Yang patut dicatat adalah *ROC AUC Score*-nya yang tertinggi, yaitu 0.9857, menunjukkan kemampuan diskriminatif superior SVM dalam membedakan secara akurat antara mahasiswa yang depresi dan tidak depresi.

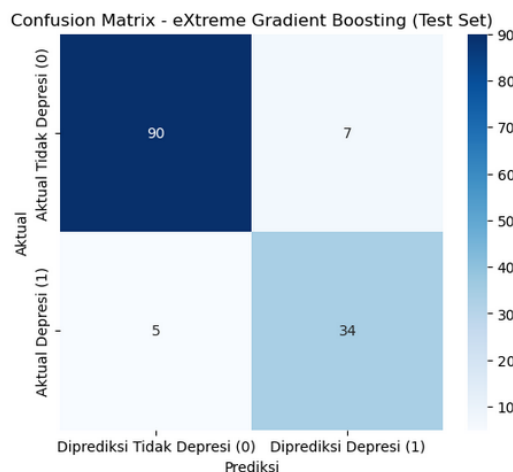


Gambar 6. *Confusion Matrix SVM*

Gambar 6 dalam model SVM menunjukkan 89 *True Negative* dan 35 *True Positive*. Terdapat 8 *False Positive* dan 4 *False Negative*, hasil *cross validation* ini, terutama kombinasi akurasi, *F1-Score*, dan *ROC AUC* yang tinggi, menjadikan SVM kandidat kuat sebagai algoritma terbaik dalam studi ini.

d. XGBoost

Pengujian model *XGBoost* setelah pelatihan pada data yang telah diskalakan dan diseimbangkan juga menghasilkan kinerja yang sangat solid. Melalui *Stratified K-Fold Cross Validation*, *XGBoost* menunjukkan Akurasi rata-rata yang sangat tinggi, yaitu 0.9404, setara dengan SVM, menandakan efisiensi dalam prediksi. Model ini meraih *Precision* sebesar 0.9351, yang merupakan nilai tertinggi dan sangat penting untuk menghindari diagnosis depresi yang keliru. *Recall*-nya tercatat 0.9483, menunjukkan kapasitas yang baik dalam mendeteksi kasus depresi yang ada. Dengan *F1-Score* 0.9409, *XGBoost* menunjukkan keseimbangan kinerja yang kuat. Meskipun *ROC AUC Score*-nya 0.9825 sedikit di bawah SVM, angka ini tetap sangat tinggi, membuktikan kemampuan diskriminatif yang unggul dalam mengklasifikasikan status depresi mahasiswa. Performa superior *XGBoost* dapat dikaitkan dengan teknik *boosting* adaptifnya yang secara iteratif memperbaiki kesalahan model sebelumnya, menghasilkan prediktor yang sangat akurat.



Gambar 7. *Confusion Matrix XGBoost*

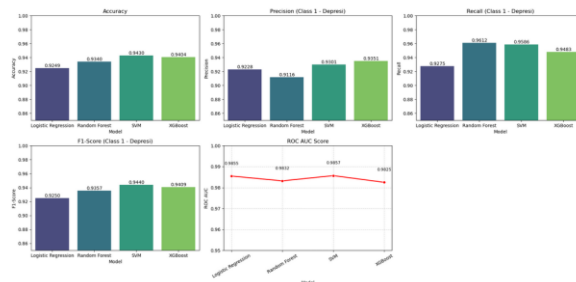
Confusion matrix pada gambar 7 menunjukkan 90 *True Negative* dan 34 *True Positives*. Terdapat 7 *False Positives* dan 5 *False Negatives*. Hasil *cross validation* menunjukkan kinerja yang sangat kuat terutama pada metrik *precision*. Jumlah *False Positives* yang relatif rendah menunjukkan bahwa *XGBoost* sangat berhati-hati dalam memberikan diagnosis depresi, mengurangi kemungkinan alarm palsu.

3.1.4 Perbandingan Kinerja Algoritma Machine Learning

Tabel 1. Perbandingan Kinerja Algoritma

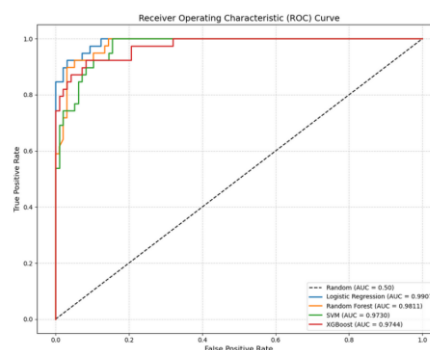
Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.9249	0.9228	0.9275	0.9250	0.9855
Random Forest	0.9340	0.9116	0.9612	0.9357	0.9832
SVM	0.9430	0.9301	0.9586	0.9440	0.9857
XGBoost	0.9404	0.9351	0.9483	0.9409	0.9825

Berdasarkan Tabel 1 dapat diamati bahwa semua model menunjukkan kinerja prediktif yang sangat tinggi dengan rata-rata akurasi di atas 92%. Secara keseluruhan, keempat algoritma ini mampu mempelajari pola dari data *PHQ-9* dan demografi untuk mengklasifikasikan status depresi mahasiswa dengan baik. Namun, terdapat perbedaan nuansa dalam metrik tertentu. Model *SVM* mencapai akurasi tertinggi sebesar 94,30%, diikuti oleh *XGBoost* 94,04%. *XGBoost* menunjukkan *precision* tertinggi 93,51% menandakan model paling efektif dalam memastikan prediksi depresi benar. *SVM* juga memiliki *precision* yang sangat baik yaitu sebesar 93,01%. *Random Forest* menjadi model dengan *recall* tertinggi yaitu 96,12% menunjukkan kemampuan yang luar biasa dalam mengidentifikasi sebagian besar mahasiswa yang benar-benar mengalami depresi. *SVM* juga menunjukkan kemampuan yang sangat baik yaitu 95,86%. *SVM* memiliki *F1-Score* tertinggi yaitu 94,40% yang mengindikasikan keseimbangan terbaik antara *precision* dan *recall* untuk kelas depresi. Nilai *F1-Score* yang tinggi ini sangat penting mengingat adanya ketidakseimbangan kelas pada *dataset*, memastikan bahwa model tidak hanya akurat tetapi juga mampu mengelola kedua jenis kesalahan klasifikasi dengan baik. Secara visual perbandingan kinerja metrik ini dapat dilihat pada gambar 8



Gambar 8. Perbandingan Kinerja Metrik

3.1.5 ROC dan AUC



Gambar 9. ROC dan AUC

Gambar 9 menunjukkan bahwa semua model memiliki kurva *ROC* yang mendekati sudut kiri atas, mengindikasikan kemampuan diskriminasi yang sangat baik. Kurva *ROC* merepresentasikan hubungan antara *True Positive Rate (Recall)* dan *False Positive Rate* pada berbagai ambang batas klasifikasi. Hal ini selaras dengan skor *ROC AUC* yang tinggi untuk semua model, yang berada di atas 0.98. Nilai *AUC* merepresentasikan probabilitas bahwa model akan memberikan peringkat yang lebih tinggi pada *instance* positif acak daripada *instance* negatif acak. *SVM* unggul dalam kemampuan diskriminasi dengan nilai *ROC AUC* tertinggi yaitu 98,57% menunjukkan performa yang superior dalam membedakan antara kelas depresi dan tidak depresi. Nilai *ROC AUC* yang sangat tinggi ini mengkonfirmasi bahwa *SVM* memiliki kemampuan terbaik dalam memisahkan kedua kelas, terlepas dari ambang batas klasifikasi yang dipilih, menjadikannya sangat handal dalam mendeteksi depresi.

3.2 Pembahasan

3.2.1 Interpretasi Hasil Evaluasi Model

Berdasarkan analisis komparatif kinerja empat algoritma *machine learning* menggunakan *Stratified K-Fold Cross Validation*, semua model menunjukkan kemampuan yang sangat baik dalam memprediksi depresi pada mahasiswa. Hal ini terbukti dengan nilai metrik evaluasi yang secara konsisten tinggi di atas 90% untuk akurasi, presisi, *recall* dan *F1-Score*, serta nilai *ROC AUC* di atas 98%. Kinerja yang solid ini menunjukkan bahwa fitur-fitur yang digunakan dari kuesioner *PHQ-9* ditambah dengan data demografi yang relevan memiliki kekuatan prediktif yang kuat dalam mengidentifikasi status depresi. Temuan ini menggarisbawahi relevansi *PHQ-9* sebagai instrumen skrining yang efektif, bahkan ketika diintegrasikan dengan algoritma *machine learning*. Keberhasilan model dalam studi ini menegaskan bahwa pendekatan berbasis AI dengan algoritma *machine learning* untuk permasalahan klasifikasi, adalah solusi yang sangat menjanjikan dalam mengotomatisasi dan meningkatkan ketepatan proses diagnostik depresi [9]. Secara praktis, ini menunjukkan bahwa sistem berbasis *machine learning* memiliki potensi besar untuk menjadi alat skrining awal yang objektif dan efisien di lingkungan pendidikan tinggi, membantu mengidentifikasi mahasiswa yang berisiko depresi lebih cepat dan skala yang lebih besar dibandingkan metode konvensional. Potensi ini sangat relevan mengingat tantangan subjektivitas dan keterbatasan sumber daya pada metode skrining tradisional.

3.2.2 Model Algoritma Spesifik untuk Prediksi Depresi

Penelitian ini berhasil mengidentifikasi bahwa *Logistic Regression*, *Random Forest*, *SVM*, dan *XGBoost* adalah model-model yang spesifik dan efektif. Keempatnya mampu mempelajari pola dari data *PHQ-9* dan demografi untuk mengklasifikasikan mahasiswa ke dalam kategori depresi atau tidak depresi dengan tingkat keberhasilan yang tinggi. Konsistensi kinerja tinggi dari model-model ini dalam konteks dataset *PHQ-9* dan demografi mahasiswa STIKes Panti Rapih Yogyakarta menegaskan relevansi dan validitas penggunaan algoritma *machine learning* untuk tujuan deteksi dini depresi, terutama dengan dukungan pra-pemrosesan data yang cermat seperti *SMOTE*. Penggunaan *SMOTE* secara signifikan berkontribusi pada kinerja model yang seimbang dengan mengatasi masalah ketidakseimbangan kelas, memastikan model tidak bias terhadap kelas mayoritas dan mampu mendeteksi kasus depresi (kelas minoritas) secara efektif. Ini juga menunjukkan bahwa karakteristik unik dari populasi mahasiswa kesehatan di STIKes Panti Rapih dapat dimodelkan secara akurat oleh algoritma *machine learning* yang dipilih. Pentingnya pemilihan algoritma yang tepat dan pra pemrosesan data yang cermat ini adalah kunci untuk menghasilkan model prediktif yang andal dan dapat diterapkan secara klinis.

3.2.3 Algoritma dengan Tingkat Akurasi Prediksi Tertinggi

Algoritma yang menunjukkan tingkat akurasi prediksi tertinggi, hasil perbandingan menunjukkan bahwa *SVM* dan *XGBoost* secara umum tampil paling unggul dalam hal akurasi dan *F1-Score* rata-rata. *SVM* secara spesifik mencapai akurasi tertinggi yaitu sebesar 94,30% dan *F1 Score* tertinggi yaitu 94,40%, serta *ROC AUC* tertinggi yaitu 98,57%. Ini mengindikasikan bahwa *SVM* adalah algoritma yang paling seimbang dan memiliki diskriminasi terbaik secara keseluruhan dalam membedakan antara kelas depresi dan tidak depresi pada dataset ini. Kinerja superior *SVM* ini dapat diatribusikan pada kemampuannya untuk menemukan *hyperplane* optimal yang memisahkan kelas dengan margin terbesar bahkan dalam data berdimensi tinggi [23]. Kesesuaian *SVM* dengan struktur data *PHQ-9* yang mungkin memiliki pola *non linear* menjadikannya pilihan yang sangat efektif. Sementara itu *Random Forest* menunjukkan kekuatan luar biasa dalam metrik *Recall* yaitu 96,12% yang sangat krusial untuk memastikan deteksi kasus depresi yang sebenarnya. Kemampuan *Random Forest* untuk mencapai *recall* tinggi ini konsisten dengan temuan studi sebelumnya [8]. Sifat *ensemble Random Forest* yang menggabungkan banyak *decision tree* memungkinkannya menangani variabilitas data dengan baik dan mengurangi *false negative* yang sangat penting dalam skrining kesehatan. *XGBoost* menunjukkan presisi tertinggi yaitu 93,51% yang berarti model ini sangat cermat dalam memprediksi depresi agar tidak salah diagnosis. Presisi yang tinggi penting untuk mencegah kecemasan yang tidak perlu akibat diagnosis depresi yang keliru [15]. Kemampuan *XGBoost* dalam memperbaiki kesalahan secara iteratif melalui *boosting* menjelaskan mengapa ia mampu mencapai tingkat presisi yang luar biasa. *Logistic Regression* menunjukkan kinerja yang solid namun sedikit dibawah ketiga model lainnya. Meskipun demikian hasil masih sangat baik, menegaskan bahwa efektivitas

prediksinya tetap signifikan, terutama mengingat kesederhanaan model ini dalam memodelkan probabilitas kejadian linier [20].

Perbedaan nuansa kinerja antar model ini menunjukkan adanya *trade-off* yang perlu dipertimbangkan berdasarkan prioritas klinis dan operasional, karena pertimbangannya untuk membangun sistem skrining yang seakurat mungkin secara keseluruhan dan seimbang dalam prediksinya maka *SVM* adalah algoritma terbaik dari ketiga algoritma lainnya.

4. KESIMPULAN

Penelitian ini berhasil mengidentifikasi bahwa *Logistic Regression*, *Random Forest*, *SVM* dan *XGBoost* merupakan model algoritma *machine learning* yang spesifik dan efektif untuk memprediksi depresi pada mahasiswa STIKes Panti Rapih Yogyakarta. Keempat algoritma ini, setelah melalui tahapan pra-pemrosesan data yang cermat termasuk penanganan ketidakseimbangan kelas dengan *SMOTE*, berhasil mempelajari pola dari data kuesioner *PHQ-9* dan informasi demografi untuk mengklasifikasikan status depresi mahasiswa dengan tingkat keberhasilan yang tinggi. Hal ini mengukuhkan potensi besar penerapan kecerdasan buatan dalam skrining kesehatan mental. Berdasarkan perbandingan metrik kinerja rata-rata dari *Stratified K-Fold Cross Validation* algoritma yang menunjukkan tingkat akurasi prediksi tertinggi bervariasi tergantung pada metrik spesifik yang menjadi prioritas. *Support Vector Machine* secara umum menunjukkan kinerja terbaik dengan nilai akurasi tertinggi 94,30%, F1-Score tertinggi 94,40%, dan *ROC AUC Score* tertinggi 98,57%. Hal ini mengindikasikan bahwa *SVM* adalah algoritma yang paling seimbang dan memiliki kemampuan diskriminasi terbaik secara keseluruhan dalam memprediksi depresi pada mahasiswa dalam dataset ini. *Random Forest* unggul dalam *Recall* tertinggi 96,12%. *XGBoost* menunjukkan presisi tertinggi 93,51% meminimalkan diagnosis depresi yang salah. Meskipun demikian, penelitian ini memiliki keterbatasan utama pada lingkup data yang spesifik, yaitu hanya bersumber dari mahasiswa STIKes Panti Rapih Yogyakarta. Hal ini dapat membatasi kemampuan generalisasi model untuk diterapkan pada populasi mahasiswa dari institusi lain atau demografi yang lebih luas. Oleh karena itu saran penelitian selanjutnya adalah perluasan dataset dengan cakupan yang lebih besar dan bervariasi dari berbagai institusi sangat disarankan untuk meningkatkan generalisasi model dan memastikan keberhasilan aplikasi model di berbagai konteks.

REFERENCES

- [1] H. Jin *et al.*, "Predicting depression using serum perfluoroalkyl and polyfluoroalkyl substances levels via interpretable machine learning," *J. Affect. Disord.*, vol. 391, pp.119976 Dec. 2025, doi: 10.1016/j.jad.2025.119976.
- [2] P. W. Simarmata and P. T. Prasetyaningrum, "Development of a Student Depression Prediction Model Based on Machine Learning with Algorithm Performance Evaluation," *J.Inf.Syst.Inform*, vol. 7, no. 2, pp. 1283–1305, Jun. 2025, doi: 10.51519/journalisi.v7i2.1087.
- [3] J. J. Mann, C. A. Michel, and R. P. Auerbach, "Improving Suicide Prevention Through Evidence-Based Strategies: A Systematic Review," *Am.J.Psychiatry*, vol.178, no 7, pp.605-608, Jul.2021, doi: 10.1176/appi.ajp.2020.20060864.
- [4] D. Arini Izzah, S. Yitnamurti, and N. M. Rehatta, "Prevalence Of Depression In First-Year Medical Students At Universitas Airlangga, Surabaya, Indonesia," *Maj. Biomorfologi*, vol. 31, no. 2, p. 39-42, Jun. 2021, doi: 10.20473/mbiom.v31i2.2021.39-43.
- [5] A. S. Ramadianto, I. Kusumadewi, F. Agiananda, and N. W. Raharjanti, "Symptoms of depression and anxiety in Indonesian medical students: association with coping strategy and resilience," *BMC Psychiatry*, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12888-022-03745-1.
- [6] N. Arif Ismail, M. Farid Adnan, and I. Fidianingsih, "Psychological Status among the Young Muslim Community in Yogyakarta Province, Indonesia, during the COVID-19 Pandemic," *IMJM*, vol. 22, no. 3, 2023, doi: 10.31436/imjm.v22i3.
- [7] D. Kurniawan, *Pengenalan Machine Learning dengan Python*. Jakarta: Gramedia, 2022.
- [8] G. H. Al Masud, R. I. Shanto, I. Sakin, and M. R. Kabir, "Effective depression detection and interpretation: Integrating machine learning, deep learning, language models, and explainable AI," *Array*, vol. 25, pp. 100375, Mar. 2025, doi: 10.1016/j.array.2025.100375.
- [9] M. Nayan *et al.*, "Comparison of the performance of machine learning-based algorithms for predicting depression and anxiety among University Students in Bangladesh: A result of the first wave of the COVID-19 pandemic," *Asian J.Soc. Health and Behav*, vol. 5, no. 2, pp. 75–84, Apr. 2022, doi: 10.4103/shb.shb_38_22.
- [10] S. M. Pahlevi, *Kecerdasan Buatan dengan Deep Learning*. Jakarta: Elex Media Komputindo, 2023.
- [11] K. Mao, Y. Wu, and J. Chen, "A systematic review on automated clinical depression diagnosis," *Nature Digit.Med*, vol.6, no 1, pp 251, Dec.2023, doi: 10.1038/s44184-023-00040-z.
- [12] Y. Luo, Y. Chen, A. Salekin, and T. Rahman, "Toward Foundation Model for Multivariate Wearable Sensing of Physiological Signals," *arXiv preprint arXiv: 2412.09758*, May 2025.
- [13] T. Vu *et al.*, "Prediction of depressive disorder using machine learning approaches: findings from the NHANES," *BMC Med. Inform. Decis. Mak.*, vol. 25, no. 1, p.27, Dec. 2025, doi: 10.1186/s12911-025-02903-1.
- [14] Y. L. Q. Song, L. Chen, H. Liu, and Y. Liu, "Machine learning algorithms to predict depression in older adults in China: a cross-sectional study," *Front. Public Health*, vol. 12, p. 1462387, 2024, doi: 10.3389/fpubh.2024.1462387.
- [15] C. Mimikou *et al.*, "Explainable Machine Learning in the Prediction of Depression," *Diagnostics*, vol. 15, no. 11, pp.1412, Jun. 2025, doi: 10.3390/diagnostics15111412.

- [16] U. M. Haque, E. Kabir, and R. Khanam, "Detection of child depression using machine learning methods," *PLoS One*, vol. 16, no. 12, pp. e0261131, Dec 2021, doi: 10.1371/journal.pone.0261131.
- [17] T. Zhu, "Analysis on the applicability of the random forest," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Aug. 2020, doi: 10.1088/1742-6596/1607/1/012123.
- [18] L. Luo *et al.*, "Predictors of depression among Chinese college students: a machine learning approach," *BMC Public Health*, vol. 25, no. 1, pp. 106, Dec. 2025, doi: 10.1186/s12889-025-21632-8.
- [19] P. Schober and T. R. Vetter, "Logistic Regression in Medical Research," *Anesthesia&Analgesia*, vol. 132, no. 2, pp. 365, Feb 2021.
- [20] A. Zaidi and A. S. M. Al Luhayb, "Two Statistical Approaches to Justify the Use of the Logistic Function in Binary Logistic Regression," *Math. Probl. Eng.*, vol. 2023, no. 1, pp. 5525675, Jan. 2023, doi: 10.1155/2023/5525675.
- [21] A. O. Hassan, I. M. Jamal, S. D. Ahmed, and A. U. Abdullahi, "Predicting Student Depression using Machine Learning: A Comparative Analysis of Machine Learning Algorithms for Early Depression Detection in Students," *Am.Int.Theism Univ.Sci.Res.J*, vol. 5, no. 1, 2025, doi: 10.63094/AITUSRJ.25.4.1.4.
- [22] V. Piccialli and M. Sciandrone, "Nonlinear optimization and support vector machines," *Ann. Oper. Res.*, vol. 314, no. 1, pp. 15–47, Jul. 2022, doi: 10.1007/s10479-022-04655-x.
- [23] N. Huda Ovirianti, M. Zarlis, and H. Mawengkang, "Support Vector Machine Using A Classification Algorithm," *J.Penelit.Tek.Inform*, vol. 6, no. 3, 2022, doi: 10.33395/sinkron.v7i3.
- [24] Q. Qiang *et al.*, "Identifying risk factors for depression and positive/negative mood changes in college students using machine learning," *Front. Public Health*, vol. 13, p. 1606947, 2025, doi: 10.3389/fpubh.2025.1606947.
- [25] P. Zhang, Y. Jia, and Y. Shang, "Research and application of XGBoost in imbalanced data," *Int. J. Distrib. Sens. Netw.*, vol. 18, no. 6, p. 1106935, Jun. 2022, doi: 10.1177/15501329221106935.
- [26] C. Yu, X. Kong, W. Yu, X. Ni, J. Chen, and X. Liao, "Machine learning models for predicting the risk of depressive symptoms in Chinese college students," *Front. Psychiatry*, vol. 16, p. 1648585, 2025, doi: 10.3389/fpsyt.2025.1648585.
- [27] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," *IEEE Access*, vol. 11, pp. 3034–3043, 2023, doi: 10.1109/ACCESS.2022.3232287.
- [28] I. Emmanuel, Y. Sun, and Z. Wang, "A machine learning-based credit risk prediction engine system using a stacked classifier and a filter-based feature selection method," *J. Big Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1186/s40537-024-00882-0.
- [29] A. Hade and M. Elhia, "Predicting mortgage credit defaults in morocco using machine learning approaches," *Discover Artif. Intell*, vol. 5, no. 1, Dec. 2025, doi: 10.1007/s44163-025-00303-y.