

PREDIKSI TINGKAT DEPRESI MAHASISWA MENGGUNAKAN MODEL HYBRID XGBOOST DAN SVM BERDASARKAN PHQ-9, IPK, TAHUN AKADEMIK, DAN USIA

Ananda Kallila^{1*}, Muhamad Bahrul Ulum²,

^{1,2}Fakultas Ilmu Komputer, Universitas Esa Unggul; Jl. Arjuna Utara No.9, Duri Kepa, Kec. Kb. Jeruk, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta 11510; (021) 39529950

Keywords:

Depresi,
XGBoost,
Support Vector Machine,
PHQ-9,
Hybrid

Correspondent Email:

suyennanda26@student.esa
unggul.ac.id

Abstrak. Kesehatan mental mahasiswa menjadi perhatian penting karena tekanan akademik dan sosial yang tinggi dapat memicu gejala depresi. Penelitian ini bertujuan membangun model klasifikasi tingkat depresi berdasarkan kuesioner PHQ-9 dan data demografis seperti usia, IPK, dan tahun akademik. Tiga pendekatan digunakan: Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), dan model hybrid. Tahapan penelitian mencakup preprocessing data (imputasi, encoding, normalisasi), pelatihan model, dan evaluasi dengan akurasi, precision, recall, F1-score, dan AUC. Model hybrid dibentuk dengan menggunakan output probabilitas XGBoost sebagai input pelatihan model SVM. Hasil evaluasi menunjukkan SVM memiliki akurasi tertinggi sebesar 96%, diikuti hybrid 93,5% dan XGBoost 93%. Namun, pada pengujian data uji individual, model hybrid menunjukkan prediksi yang lebih tepat dibanding SVM. Seluruh model diimplementasikan dalam antarmuka berbasis web agar hasil klasifikasi mudah diakses oleh pengguna. Penelitian ini menunjukkan bahwa baik model tunggal maupun hybrid memiliki potensi besar dalam mengklasifikasikan tingkat depresi secara non-diagnostik. SVM unggul dalam klasifikasi makro, XGBoost efektif untuk pola fitur kompleks, dan pendekatan hybrid menawarkan solusi prediksi yang lebih stabil dan adaptif terhadap variasi data mahasiswa.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. Mental health among students has become a major concern, as academic and social pressures can trigger symptoms of depression. This study aims to develop a classification model for predicting depression levels based on PHQ-9 questionnaire responses and demographic data such as age, GPA, and academic year. Three approaches were applied: Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and a hybrid model. The research stages included data preprocessing (imputation, encoding, normalization), model training, evaluation using accuracy, precision, recall, F1-score, and AUC metrics. The hybrid model was constructed by feeding XGBoost's predicted probabilities as input features into the SVM model. Evaluation results show that SVM achieved the highest accuracy at 96%, followed by the hybrid model at 93,5%, and XGBoost at 93%. However, in individual test sample predictions, the hybrid model demonstrated higher precision than SVM. All models were deployed through a web based interface to allow users to access the predictions interactively. This study highlights that both individual and hybrid models hold strong potential in non diagnostic classification of depression levels. SVM excels in overall classification, XGBoost is effective for capturing complex feature patterns, and the hybrid

approach provides a more stable and adaptive prediction in diverse student data contexts.

1. PENDAHULUAN

Masalah kesehatan mental di kalangan mahasiswa telah menjadi perhatian yang signifikan, didorong oleh tekanan akademis, sosial, dan finansial yang dapat memicu gejala depresi [1]. Depresi, yang ditandai dengan kesedihan yang terus-menerus, kehilangan minat, gangguan tidur, perubahan nafsu makan, dan kelelahan, tidak hanya memengaruhi kesejahteraan emosional tetapi juga mengurangi fungsi akademis dan sosial [2]. Menurut Organisasi Kesehatan Dunia (WHO), depresi adalah salah satu penyebab utama kecacatan global, dengan prevalensi yang meningkat di kalangan dewasa muda, termasuk mahasiswa [1], [2]. Di Indonesia, Survei Kesehatan Mental Remaja Nasional Indonesia (I-NAMHS) melaporkan bahwa 34,9% remaja mengalami masalah kesehatan mental dalam setahun terakhir, sementara hanya 2,6% yang mengakses layanan konseling profesional [3]. Hal ini mencerminkan kesenjangan pengobatan yang substansial, diperburuk oleh stigma, akses terbatas ke layanan profesional, dan kecenderungan mahasiswa untuk mengandalkan diagnosis sendiri melalui sumber non-profesional [3], [4], [5].

Stigma sosial tetap menjadi penghalang utama bagi mahasiswa dalam mencari dukungan kesehatan mental, karena banyak yang merasa malu atau takut akan penilaian negatif dari orang lain [4]. Tantangan logistik seperti biaya tinggi dan ketersediaan layanan kesehatan mental yang terbatas juga menghambat akses ke perawatan [4]. Selain itu, banyak siswa melakukan diagnosis sendiri menggunakan informasi dari internet, yang menimbulkan potensi risiko salah diagnosis dan memperburuk kondisi mereka [5].

Studi sebelumnya telah menunjukkan bahwa faktor internal dan eksternal, khususnya lingkungan akademik dan sosial, memengaruhi kesehatan mental siswa [2], [6]. Algoritma pembelajaran mesin seperti Support Vector Machine (SVM) dan Extreme Gradient Boosting (XGBoost) telah banyak diterapkan

untuk prediksi kesehatan mental, menunjukkan akurasi yang menjanjikan dalam mendeteksi gangguan depresi dari data survei dan klinis [6], [7], [8]. SVM efektif dalam menangani data berdimensi tinggi dan membangun batas keputusan optimal [6], [7], sedangkan XGBoost unggul dalam memodelkan interaksi fitur yang kompleks dan mengelola dataset yang tidak seimbang [8]. Namun, setiap model memiliki keterbatasannya—SVM tidak secara inheren memberikan output probabilistik, dan XGBoost dapat sensitif terhadap ketidakseimbangan kelas dan overfitting tanpa penyetelan yang tepat [8].

Untuk mengatasi keterbatasan ini, penelitian ini mengusulkan pendekatan pemodelan hibrida yang menggunakan output probabilistik XGBoost sebagai fitur input untuk SVM. Pendekatan ini bertujuan untuk memanfaatkan kekuatan XGBoost dalam ekstraksi fitur dan kemampuan SVM dalam pemisahan kelas optimal, yang berpotensi meningkatkan stabilitas dan kinerja prediksi untuk kelas minoritas. Penelitian ini berfokus pada pengembangan model prediktif untuk tingkat depresi hanya menggunakan respons dari Kuesioner Kesehatan Pasien-9 (PHQ-9) [9], tanpa memasukkan variabel demografis seperti usia, IPK, atau tahun akademik, untuk menjaga kesederhanaan, memastikan privasi, dan fokus pada indikator psikologis.

Tujuan penelitian ini adalah untuk (1) mengembangkan sistem klasifikasi yang akurat untuk tingkat depresi mahasiswa berdasarkan respons PHQ-9, (2) mengimplementasikan dan membandingkan SVM, XGBoost, dan model hibrida yang diusulkan, dan (3) menerapkan model dalam antarmuka berbasis web untuk prediksi yang mudah diakses dan ramah pengguna. Data dikumpulkan melalui kuesioner PHQ-9 daring yang didistribusikan kepada mahasiswa aktif, diikuti oleh langkah-langkah pra-pemrosesan termasuk penanganan nilai yang hilang, pengkodean, normalisasi, dan pemilihan fitur. Model dilatih dan dievaluasi menggunakan metrik akurasi, presisi, recall, F1-score, dan Area Under the Curve (AUC).

Model hibrida dirancang sebagai pipeline dua tahap, di mana XGBoost menghasilkan kelas vektor probabilitas yang kemudian diklasifikasikan oleh SVM.

Dengan menggabungkan teknik pembelajaran mesin yang kuat dengan sistem berbasis web yang mudah diakses dan non-diagnostik, studi ini bertujuan untuk berkontribusi pada deteksi dini dan kesadaran akan depresi di kalangan mahasiswa, yang pada akhirnya mendukung inisiatif kesehatan mental di lembaga pendidikan tinggi.

2. TINJAUAN PUSTAKA

2.1. Kesehatan Mental

Kesehatan mental adalah kondisi kesejahteraan individu yang memungkinkan seseorang mampu menghadapi tekanan hidup, menjalin hubungan sosial yang sehat, bekerja secara produktif, dan berkontribusi kepada masyarakat. Menurut Organisasi Kesehatan Dunia atau World Health Organization, kesehatan mental bukan sekadar tidak adanya gangguan mental, tetapi mencakup keseimbangan emosional, psikologis, dan sosial. Gangguan kesehatan mental dapat berdampak pada fungsi kognitif, emosional, dan sosial individu serta sering kali tidak terdeteksi sejak dini.

2.2. Depresi

Depresi merupakan gangguan suasana hati yang ditandai dengan perasaan sedih berkepanjangan, kehilangan minat, gangguan tidur dan nafsu makan, kelelahan, serta penurunan konsentrasi hingga munculnya pikiran untuk menyakiti diri sendiri. Penyebab depresi bersifat multifaktorial, meliputi faktor biologis, psikologis, dan lingkungan. Berdasarkan laporan World Health Organization, depresi termasuk penyebab utama disabilitas global dan banyak terjadi pada remaja serta dewasa muda, termasuk mahasiswa yang menghadapi tekanan akademik dan sosial.

2.3. Kuesioner (PHQ-9)

Kuesioner merupakan instrumen penelitian untuk mengumpulkan data secara sistematis dan harus memenuhi aspek validitas serta reliabilitas. Dalam penelitian ini digunakan Patient Health Questionnaire-9 (PHQ-9), yaitu

instrumen self-report yang terdiri dari sembilan pertanyaan berbasis skala Likert 0–3 untuk menilai gejala depresi selama dua minggu terakhir. Skor total berkisar 0–27 dan diklasifikasikan menjadi empat kategori tingkat depresi. PHQ-9 digunakan sebagai sumber fitur numerik dalam proses klasifikasi machine learning.

2.4. Machine Learning

Machine Learning adalah cabang kecerdasan buatan yang memungkinkan sistem belajar dari data untuk melakukan prediksi atau pengambilan keputusan tanpa pemrograman eksplisit. Prosesnya meliputi tahap pelatihan (training) untuk membangun model dan pengujian (testing) untuk mengevaluasi performa. Berdasarkan metode pembelajarannya, machine learning terbagi menjadi supervised learning, unsupervised learning, dan reinforcement learning, yang masing-masing digunakan untuk tugas klasifikasi, pengelompokan, maupun optimasi berbasis reward.

2.5. Deep Learning

Deep Learning merupakan bagian dari machine learning yang menggunakan jaringan saraf tiruan berlapis (neural networks) untuk memproses data kompleks seperti gambar, teks, dan suara. Melalui struktur berlapis, model dapat mengekstraksi fitur secara hierarkis dari data mentah hingga pola yang lebih kompleks. Pendekatan ini banyak digunakan dalam pengolahan citra, pengenalan suara, dan pemrosesan bahasa alami karena kemampuannya dalam menghasilkan representasi fitur yang mendalam.

2.6. XGBoost

XGBoost adalah algoritma berbasis ensemble yang menggunakan metode Gradient Boosting Decision Trees (GBDT) untuk meningkatkan akurasi prediksi dalam klasifikasi dan regresi. Algoritma ini bekerja dengan menggabungkan beberapa model lemah secara iteratif untuk meminimalkan kesalahan prediksi. Keunggulan XGBoost terletak pada kecepatan komputasi, kemampuan menangani data kompleks, serta fitur regularization yang membantu mencegah overfitting dan meningkatkan generalisasi model.[10]

2.7. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma klasifikasi dan regresi yang bekerja dengan mencari hyperplane optimal untuk memisahkan dua kelas data dengan margin maksimum. Pendekatan ini efektif dalam menangani data berdimensi tinggi dan mampu meningkatkan kemampuan generalisasi model terhadap data baru. Dengan prinsip memaksimalkan margin, SVM cenderung lebih tahan terhadap overfitting dibandingkan beberapa metode lainnya [11].

2.8. Streamlit

Streamlit adalah framework berbasis Python yang digunakan untuk membangun aplikasi web interaktif secara cepat tanpa memerlukan keahlian front-end. Dengan pendekatan deklaratif, pengembang dapat mengintegrasikan model machine learning ke dalam antarmuka web hanya dengan beberapa baris kode. Streamlit mendukung visualisasi data, form input, dan deployment berbasis cloud, sehingga memudahkan penyajian hasil analisis kepada pengguna akhir.

2.9. Python

Python adalah bahasa pemrograman tingkat tinggi yang populer karena sintaksnya yang sederhana, fleksibel, dan mendukung berbagai paradigma pemrograman. Python banyak digunakan dalam pengembangan perangkat lunak, analisis data, dan kecerdasan buatan karena memiliki pustaka yang lengkap serta komunitas yang besar. Keunggulan ini menjadikan Python sebagai bahasa utama dalam pengembangan sistem machine learning.

2.10. TensorFlow

TensorFlow adalah pustaka open-source yang dikembangkan oleh Google untuk membangun dan melatih model machine learning serta deep learning secara efisien. TensorFlow mendukung komputasi pada CPU, GPU, dan TPU, sehingga mampu menangani pemrosesan data skala besar. Integrasinya dengan Keras mempermudah pengembangan model neural network dalam berbagai aplikasi seperti pengenalan citra dan prediksi berbasis data.

2.11. Keras

Keras adalah pustaka tingkat tinggi untuk membangun dan melatih model deep learning

secara cepat dan modular. Keras dirancang agar mudah digunakan, memungkinkan pembuatan model jaringan saraf hanya dengan beberapa baris kode. Dengan antarmuka yang intuitif dan integrasi yang baik dengan TensorFlow, Keras banyak digunakan dalam pengembangan model pembelajaran mesin berbasis neural network.

3. METODE PENELITIAN

A. Perumusan Masalah

Tujuan utama penelitian ini adalah untuk mengembangkan model klasifikasi multikelas yang mampu memprediksi tingkat depresi mahasiswa berdasarkan respons terhadap kuesioner PHQ-9. Model klasifikasi dirancang untuk memetakan fitur ke label kelas, yang mewakili enam tingkat keparahan depresi. Instrumen PHQ-9 telah divalidasi untuk mengukur depresi pada populasi mahasiswa [9], sementara algoritma pembelajaran mesin seperti XGBoost dan Support Vector Machine (SVM) telah menunjukkan kinerja yang kuat dalam deteksi gangguan kesehatan mental dari data berbasis survei [12], [13].

B. Dataset

Dataset untuk penelitian ini dikumpulkan dari mahasiswa aktif yang menggunakan kuesioner PHQ-9, yang terdiri dari sembilan item yang diberi skor pada skala Likert dari 0 ("Tidak sama sekali") hingga 3 ("Hampir setiap hari") [14]. PHQ-9 telah menunjukkan konsistensi internal yang tinggi ($\alpha > 0,8$) dan reliabilitas untuk mengidentifikasi tingkat keparahan depresi berdasarkan total skor gejala [9], [15]

Berdasarkan total skor PHQ-9, respons dikategorikan ke dalam empat tingkat keparahan depresi: Tidak Depresi, Depresi Ringan, Depresi Sedang, dan Depresi Berat. Tabel 1 merangkum rentang skor PHQ-9, jumlah sampel, dan proporsi untuk setiap kategori.

Tabel 1. Kelas Depresi

Class	PHQ-9 Score	Samples	%
No Depression	0–5	330	33.0
Mild Depression	6–11	314	31.4
Moderate Depression	12-17	269	26.9
Severe Depression	>17	87	8.7

C. Praproses Data

1) Pembersihan Data

Tahap awal melibatkan penanganan nilai yang hilang dan memastikan konsistensi data [16]. Fitur numerik dengan nilai yang hilang diimputasi menggunakan rata-rata, sedangkan fitur kategorikal diimputasi menggunakan modus. Pencilan pada variabel numerik dideteksi menggunakan metode Rentang Interkuartil (IQR) dan ditangani melalui pemotongan untuk menghindari distorsi dalam analisis. Respons PHQ-9 dikonversi dari teks kategorikal ke nilai numerik ordinal (0–3). Respons yang tidak valid di luar rentang yang valid dihapus atau diimputasi. Entri duplikat dihapus untuk mencegah bias. Dalam kasus di mana satu atau dua skor item PHQ-9 hilang, imputasi dilakukan menggunakan nilai rata-rata item yang sesuai di seluruh dataset; entri dengan lebih dari dua skor yang hilang dihapus.

2) Normalisasi

Untuk memastikan semua fitur berkontribusi secara sama pada model, normalisasi data diterapkan [17], [18]. Skala Standar digunakan untuk mengubah nilai numerik sehingga setiap fitur memiliki rata-rata nol dan deviasi standar satu. Proses ini diterapkan pada semua fitur numerik kontinu dan skor PHQ-9 setelah pengkodean kategorikal.

3) Pemilihan Fitur

Pemilihan fitur dilakukan menggunakan teknik kepentingan fitur XGBoost [19]. Pendekatan ini menghitung kontribusi setiap fitur terhadap

kinerja model, memungkinkan penghapusan fitur yang kurang relevan untuk meningkatkan efisiensi komputasi dan mengurangi risiko overfitting.

4) Pengkodean Kategorikal

Variabel kategorikal, termasuk respons PHQ-9, dikodekan ke dalam bentuk numerik menggunakan Pengkodean Label [19]. Setiap kategori unik diberi bilangan bulat unik, memungkinkan kompatibilitas dengan algoritma pembelajaran mesin.

5) Pemisahan Data

Dataset dibagi menjadi set pelatihan dan pengujian menggunakan rasio 80:20 [20]. Pengambilan sampel bertingkat digunakan untuk mempertahankan proporsi distribusi kelas di kedua set. Ini memastikan bahwa set pengujian secara akurat mewakili keseluruhan dataset dan bahwa ketidakseimbangan kelas tidak membiaskan evaluasi.

D. Pengembangan Model

1) XGBoost

Selama fase pelatihan, model XGBoost digunakan untuk memprediksi tingkat depresi siswa berdasarkan skor PHQ-9 yang telah diproses. XGBoost dipilih karena kemampuannya untuk menangani dataset dengan fitur yang relatif sedikit namun hubungan non-linear yang kompleks, serta kemampuannya untuk mengoptimalkan kinerja model melalui pendekatan gradient boosting [21], [22].

Pelatihan dimulai dengan menetapkan parameter awal XGBoost. Parameter `n_estimators` ditetapkan menjadi 100, yang berarti model secara berurutan membangun hingga 100 pohon keputusan, di mana setiap pohon bertujuan untuk memperbaiki kesalahan prediksi dari pohon sebelumnya. Parameter `max_depth` ditetapkan pada 3 untuk mengontrol kompleksitas model, mencegah overfitting [21], [10], dan memastikan model fokus pada pembelajaran pola-pola penting dalam data tanpa terlalu beradaptasi dengan set pelatihan. Metrik evaluasi yang digunakan adalah `mlogloss` (multiclass logarithmic loss), yang mengukur perbedaan antara distribusi probabilitas yang diprediksi dan label aktual.

XGBoost melatih model menggunakan informasi gradien dan hessian pada setiap iterasi untuk menentukan titik pemisah terbaik untuk setiap pohon [22]. Setiap pohon yang baru ditambahkan berfokus pada kesalahan residual dari prediksi sebelumnya.

Setelah semua iterasi selesai, prediksi akhir adalah kontribusi kumulatif dari semua pohon yang dibangun oleh model. Pendekatan iteratif ini menghasilkan model yang lebih stabil dan akurat dengan secara progresif memperbaiki kesalahan pada setiap tahap konstruksi pohon keputusan.

2) Support Vector Machine

Support Vector Machine (SVM) adalah algoritma pembelajaran terawasi yang banyak digunakan untuk tugas klasifikasi dan regresi. Tujuan utamanya adalah untuk mengidentifikasi hyperplane optimal yang memisahkan titik data dari kelas yang berbeda dalam ruang fitur berdimensi tinggi [23]. SVM berfokus pada memaksimalkan margin antar kelas sambil meminimalkan kesalahan klasifikasi, bertujuan untuk mencapai batas keputusan yang dapat digeneralisasi dengan baik ke data yang belum pernah dilihat sebelumnya.

Proses pelatihan dimulai dengan data input berlabel, di mana setiap instance terdiri dari vektor fitur (input) dan label yang sesuai (output). SVM mencari hyperplane yang membagi ruang fitur menjadi wilayah-wilayah berbeda untuk setiap kelas, dengan margin terbesar di antara mereka. Margin mengacu pada jarak antara hyperplane dan titik data terdekat dari setiap kelas, yang dikenal sebagai vektor pendukung. Vektor pendukung ini memiliki pengaruh terbesar dalam menentukan posisi dan orientasi hyperplane.

Dalam penelitian ini, model SVM dilatih dengan parameter regularisasi (C) untuk mengontrol keseimbangan antara memaksimalkan margin dan meminimalkan kesalahan klasifikasi. Nilai C yang lebih tinggi bertujuan untuk mengklasifikasikan semua contoh pelatihan dengan benar, yang berpotensi menyebabkan overfitting, sementara nilai yang lebih rendah memungkinkan margin yang lebih

lunak yang dapat meningkatkan generalisasi. Fungsi kernel juga didefinisikan untuk memetakan ruang input ke ruang fitur berdimensi lebih tinggi, memungkinkan SVM untuk menangani batas klasifikasi non-linier [24],[11].

Dengan mengoptimalkan parameter ini, SVM mampu mempelajari batas kelas yang kompleks sambil mempertahankan ketahanan terhadap overfitting, terutama saat bekerja dengan dataset berukuran kecil hingga menengah.

3) Hibrida

Pengembangan model hibrida dalam penelitian ini dimulai dengan melatih model Extreme Gradient Boosting (XGBoost) menggunakan data yang telah melalui pra-pemrosesan. XGBoost membangun serangkaian pohon keputusan secara iteratif (boosting), di mana setiap pohon dilatih untuk memperbaiki kesalahan prediksi yang dibuat oleh pohon sebelumnya. Setelah proses pelatihan, model menghasilkan output probabilitas kelas untuk setiap sampel pelatihan dan pengujian, yang mewakili kepercayaan model dalam menetapkan sampel tertentu ke setiap kategori tingkat depresi.

Probabilitas kelas yang dihasilkan oleh XGBoost ini tidak digunakan secara langsung sebagai output akhir tetapi malah dimasukkan sebagai fitur input ke model Support Vector Machine (SVM). Pada tahap ini, SVM dilatih menggunakan data baru yang direpresentasikan sebagai vektor probabilitas dari XGBoost. Oleh karena itu, SVM tidak beroperasi langsung pada fitur asli (seperti skor PHQ-9) tetapi pada ruang fitur yang diperkaya yang dihasilkan oleh XGBoost. SVM kemudian membangun hyperplane di ruang ini untuk memisahkan kelas depresi berdasarkan pola yang diringkas oleh model boosting.

Pendekatan ini lebih tepat dikategorikan sebagai pembelajaran sekuensial daripada generalisasi bertumpuk. Dalam struktur ini, XGBoost berfungsi sebagai generator fitur tingkat lanjut, sementara SVM berfungsi sebagai pengklasifikasi akhir yang mempelajari pola dari representasi probabilistik. Hasil klasifikasi akhir dihasilkan oleh SVM, berdasarkan hubungan antar probabilitas kelas

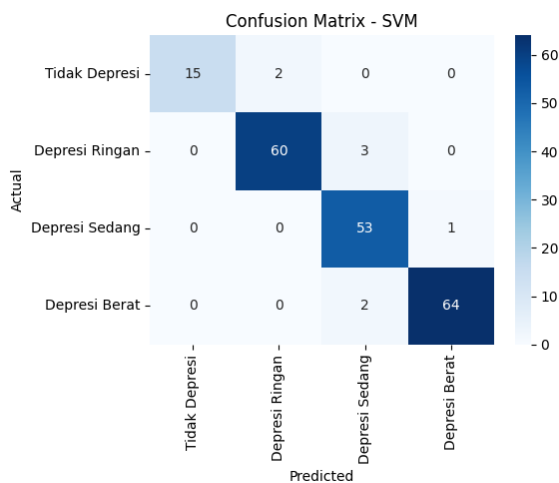
yang sebelumnya dipelajari oleh model boosting.

4. HASIL DAN PEMBAHASAN

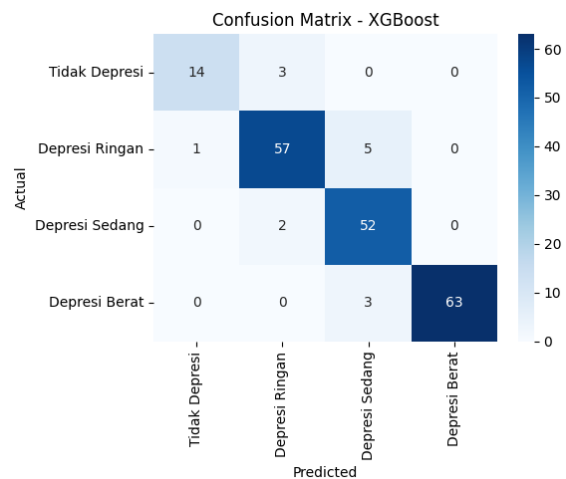
Evaluasi model dilakukan menggunakan matriks kebingungan, akurasi, dan kurva ROC–AUC pada tiga pendekatan: XGBoost, SVM, dan model hibrida. Ketiga metrik evaluasi ini dipilih untuk memberikan ukuran kinerja secara keseluruhan dan pandangan rinci tentang prediksi spesifik kelas.

A. Matriks Kebingungan

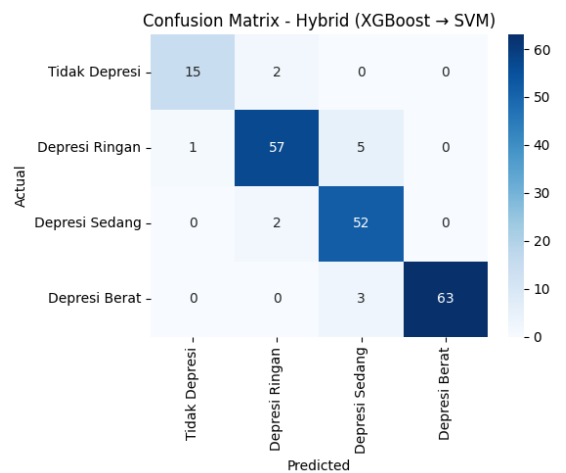
Matriks kebingungan XGBoost, SVM, dan Hibrida menunjukkan bahwa ketiga model tersebut mampu mengklasifikasikan tingkat depresi secara efektif, meskipun dengan sedikit perbedaan dalam kesalahan klasifikasi. XGBoost cenderung berkinerja lebih baik pada kelas mayoritas, sementara SVM menunjukkan hasil yang lebih stabil di semua kategori. Model hibrida mengurangi beberapa kesalahan klasifikasi yang terlihat pada model dasar, menghasilkan prediksi yang lebih seimbang di keempat tingkat depresi.



Gambar 1. Confusion Matrix SVM



Gambar 2. Confusion Matrix XGBoost

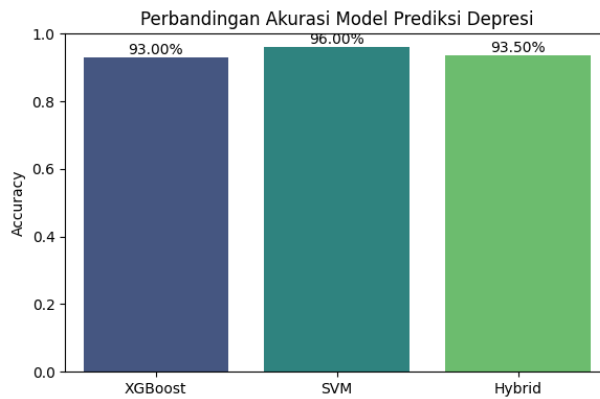


Gambar 3. Confusion Matrix Hybrid

Model SVM menunjukkan hasil yang paling konsisten, terutama pada kelas Depresi Berat, dengan 64 prediksi yang benar dan hanya 2 kesalahan klasifikasi. Model XGBoost, meskipun juga berkinerja kuat, menunjukkan sedikit lebih banyak kesalahan pada kelas Tanpa Depresi dan Depresi Ringan. Model Hibrida menawarkan hasil yang seimbang, dengan kinerja yang mendekati SVM, tetapi masih menunjukkan kesalahan klasifikasi kecil pada kategori Depresi Ringan dan Sedang. Temuan ini menunjukkan bahwa SVM lebih tepat dalam memisahkan kelas, sementara model XGBoost dan Hibrida memberikan alternatif yang kompetitif dengan kemampuan generalisasi yang baik.

B. Matriks Evaluasi

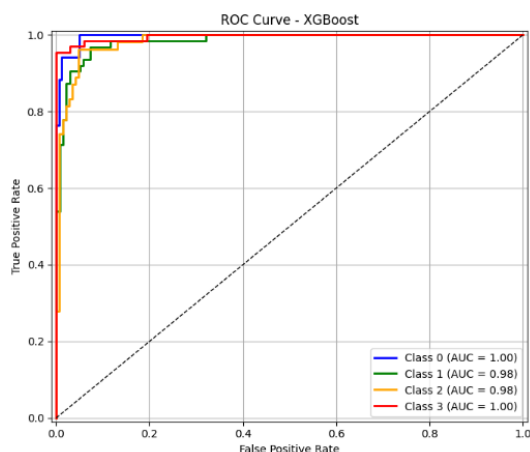
Perbandingan akurasi mengungkapkan bahwa model SVM mencapai akurasi tertinggi sebesar 96,00%, melampaui model Hibrida (93,50%) dan XGBoost (93,00%).



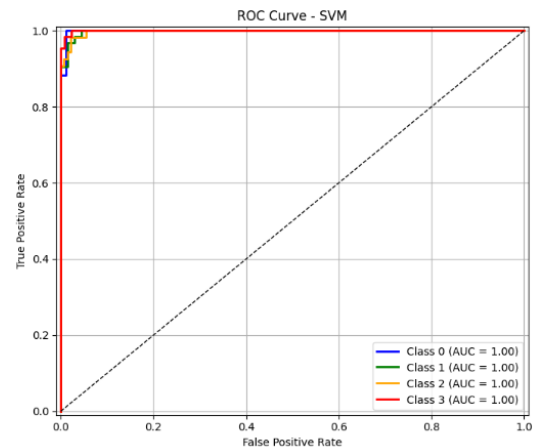
Gambar 4. Perbandingan Akurasi Model

Hasil ini menunjukkan bahwa SVM memiliki kemampuan terkuat untuk menangkap batas keputusan antara kategori depresi. Namun, model hibrida masih berkinerja kompetitif dengan memanfaatkan output probabilistik dari XGBoost sebagai fitur input lanjutan untuk SVM. Meskipun akurasinya sedikit di bawah SVM, pendekatan hibrida memberikan perspektif tambahan yang menekankan klasifikasi seimbang di berbagai tingkat depresi.

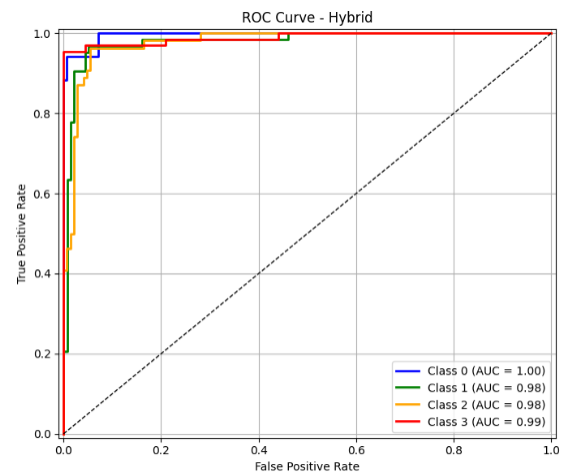
C. Kurva ROC dan AUC



Gambar 5. ROC Curve XGBoost



Gambar 6. ROC Curve SVM



Gambar 7. ROC Curve Hybrid

Analisis ROC–AUC lebih lanjut memvalidasi kinerja ketiga model tersebut. Model SVM mencapai nilai AUC sempurna sebesar 1,00 di keempat kelas, yang mencerminkan kemampuan diskriminatif yang sangat baik dalam membedakan antara kategori depresi. Sementara itu, XGBoost dan model Hibrida menghasilkan nilai AUC berkisar antara 0,98 dan 1,00, yang juga menunjukkan keandalan klasifikasi yang kuat dalam prediksi multikelas. Hasil ini menunjukkan bahwa semua model sangat efektif, dengan SVM menunjukkan stabilitas tertinggi dan model hibrida mempertahankan kinerja yang kompetitif dengan menggabungkan representasi fitur yang dihasilkan oleh XGBoost.

5. KESIMPULAN

Studi ini berhasil mengembangkan sistem yang mampu mendeteksi tingkat depresi secara akurat di kalangan mahasiswa dengan memanfaatkan data kuesioner PHQ-9 sebagai fitur utama, tanpa bergantung pada atribut demografis tambahan. Pendekatan pembelajaran mesin terbukti efektif dalam mengklasifikasikan depresi ke dalam empat kategori: Tidak Depresi, Depresi Ringan, Depresi Sedang, dan Depresi Berat.

Tiga model klasifikasi—Support Vector Machine (SVM), XGBoost, dan model Hibrida—diimplementasikan dan dibandingkan dalam hal kinerja. XGBoost mencapai akurasi 93%, SVM memperoleh akurasi tertinggi 96%, dan model Hibrida mencapai 94%. Pendekatan Hibrida dibangun dengan menggunakan output probabilistik XGBoost sebagai fitur input untuk SVM, memungkinkan kombinasi kekuatan kedua algoritma dan meningkatkan ketahanan klasifikasi multikelas.

Selain itu, hasil prediksi diintegrasikan ke dalam antarmuka berbasis web interaktif yang dirancang khusus untuk mahasiswa. Sistem ini memungkinkan pengguna untuk memilih model klasifikasi, mengisi kuesioner PHQ-9, dan menerima prediksi waktu nyata mengenai tingkat depresi mereka. Hasilnya disajikan dengan visualisasi yang jelas dan label informatif, memberikan pengguna pemahaman yang mudah diakses dan objektif tentang kondisi kesehatan mental mereka.

DAFTAR PUSTAKA

- [1] C. F. P. S. A. Daza Vergaray, J. C. H. Miranda, J. B. Cornelio, A. R. López Carranza, "Predicting depression in university students using stacking ensemble techniques over oversampling method," *Informatics Med. Unlocked*, vol. 41, 2023, doi: doi: 10.1016/j.imu.2023.101295.
- [2] N. I. F. Kaligis, R. Indrayanti, "Mental health problems and needs among transitional age youth in Indonesia," *Int. J. Environ. Res. Public Health*, vol. 18, no. 6, 2021, doi: doi: 10.3390/ijerph18062912.
- [3] R. C. A. A. Khan, O. Chaudhari, "A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation," *Expert Syst. Appl.*, vol. 244, 2024, doi: doi: 10.1016/j.eswa.2023.122778.
- [4] S. C. B. R. Krebs, S. S. Bagui, D. Mink, "Applying multi-class support vector machines: One-vs.-One vs. One-vs.-All on the UWF-ZeekDataFall22 dataset," *Electronics*, vol. 13, no. 19, 2024, doi: doi: 10.3390/electronics13193916.
- [5] A. M. S. Dalal, E. M. Onyema, "Hybrid XGBoost model with hyperparameter tuning for prediction of liver disease with better accuracy," *World J. Gastroenterol.*, vol. 28, no. 46, pp. 6551–6563, 2022, doi: doi: 10.3748/wjg.v28.i46.6551.
- [6] Y. S. P. Zhang, Y. Jia, "Research and application of XGBoost in imbalanced data," *Int. J. Distrib. Sens. Networks*, vol. 18, no. 6, 2022, doi: doi: 10.1177/15501329221106935.
- [7] H. H. and A. Idri, "A comparative study on the impact of categorical encoding on black box model interpretability," *Proc. 13th Int. Conf. Data Sci. Technol. Appl. (DATA 2024)*, pp. 384–391, 2024, doi: 384–391, doi: 10.5220/0012766300003756.
- [8] H. R. A. M. Razavi, S. T. Yazdani, H. R. Rabiee, "Machine learning, deep learning and data preprocessing techniques for detection, prediction, and monitoring of stress and stress-related mental disorders: A scoping review," *JMIR Ment. Heal.*, vol. 11, pp. 1–28, 2023, doi: doi: 10.2196/53714.
- [9] O. T. T. Emmanuel, T. Maupong, D. Mpoeleng, T. Semong, B. Mphago, "A survey on missing data in machine learning," *J. Big Data*, vol. 8, no. 1, 2021, doi: doi: 10.1186/s40537-021-00516-9.
- [10] S. Schwartz, S. Ross, J. Bryant, and J. Duncan, "Mental and physical health among students at a private university that held in-person classes during the COVID-19 pandemic," *J. Am. Coll. Heal.*, vol. 72, pp. 1–11, Mar. 2022, doi: doi: 10.1080/07448481.2022.2052074.
- [11] R. A. Saputra, "ANALISIS SENTIMEN REVIEW SKINCARE SKINTIFIC DENGAN ALGORITMA SUPPORT VECTOR MACHINE," vol. 12, no. 2, 2024.
- [12] J. B. P. A. Perez-Lebel, G. Varoquaux, M. Le Morvan, J. Josse, "Benchmarking missing-values approaches for predictive models on health databases," *Gigascience*, vol. 11, pp. 1–22, 2022, doi: doi: doi: doi: 10.1093/gigascience/giab000.

- 10.1093/gigascience/giac013.
- [13] J. M. M. A. Rahman, T. A. Dhira, A. R. Sarker, "Validity and reliability of the Patient Health Questionnaire scale (PHQ-9) among university students of Bangladesh," *PLoS One*, vol. 17, no. 6, pp. 1–18, 2022, doi: doi: 10.1371/journal.pone.0269634.
- [14] A. C. D.-B. C. A. Cassiani-Miranda, O. Scopetta, M. A. Barrios-Villadiego, A. F. Tirado-Otálvaro, "Construct validity of the PHQ-9 in university students in Colombia: A Rasch analysis approach," *Med. Clínica y Soc.*, vol. 9, no. 1, 2025, doi: doi: 10.52379/mcs.v9.521.
- [15] X. Q. and X. Huang, "Machine learning-driven identification of key risk factors for predicting depression among nurses," *BMC Nurs.*, vol. 24, no. 1, 2025, doi: doi: 10.1186/s12912-025-02957-6.
- [16] U. A. and P. Agarwal, "DepXGBoot: Depression detection using a robust tuned extreme gradient boosting model generator," *IAES Int. J. Artif. Intell.*, vol. 13, no. 4, pp. 4352–4363, 2024, doi: doi: 10.11591/ijai.v13.i4.pp4352-4363.
- [17] N. D. R. Pranekeviciene, E. T. Leskauskas, M. Adomaitiene, "Validation of the PHQ-9 and GAD-7 in primary care and psychiatry settings," *BMC Psychiatry*, vol. 22, no. 1, pp. 1–10, 2022, doi: doi: 10.1186/s12888-022-04452-5.
- [18] H. A. R. Ibrahim, A. Rahman, M. H. Ali, "Extreme Gradient Boosting-based prediction for depression among university students," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 5, pp. 725–732, 2023, doi: doi: 10.14569/IJACSA.2023.0140588.
- [19] T. D. N. N. Vu, A. Q. Nguyen, "Prediction of depressive disorder using machine learning approaches," *IEEE Access*, vol. 13, pp. 11987–11999, 2025, doi: doi: 10.1109/ACCESS.2025.3500142.
- [20] R. K. T. S. Singh, R. K. Saini, A. Sharma, "Machine learning algorithms for detecting mental stress: A comprehensive review," *Artif. Intell. Med.*, vol. 144, p. 102658, 2024, doi: doi: 10.1016/j.artmed.2024.102658.
- [21] K. K. R. Indonesia, "Survei Nasional Kesehatan Mental Remaja Indonesia (I-NAMHS): Laporan Nasional," Kementerian Kesehatan RI.
- [22] W. H. Organization, "World Health Organization, 'Depressive disorder (depression),'" [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/depression>
- [23] O. V. Chukwuma *et al.*, "A Systematic Review of the Effect of Stigmatization on Psychiatric Illness Outcomes.," *Cureus*, vol. 16, no. 6, p. e62642, Jun. 2024, doi: 10.7759/cureus.62642.
- [24] L. Guo *et al.*, "Prevalence and changes in depressive symptoms among postgraduate students: A systematic review and meta-analysis from 1980 to 2020.," *Stress Heal. J. Int. Soc. Investig. Stress*, vol. 37, no. 5, pp. 835–847, Dec. 2021, doi: 10.1002/smi.3045.