

# Identifikasi Faktor Dominan Risiko Depresi Remaja Menggunakan XGBoost dengan Interpretasi SHAP

Hartati Tammamah Lubis<sup>1\*</sup>, Randy Brilliant Chandra<sup>2</sup>, Muhammad Nasri Gea<sup>3</sup>, Erica Rian Safitri<sup>4</sup>

<sup>1</sup>Sistem Informasi, Sains dan Teknologi, Universitas Prima Indonesia, Medan, Indonesia

<sup>2</sup>Sistem Informasi, Institut Bisnis IT&B, Medan, Indonesia

<sup>3</sup>Manajemen Informatika, Akademi Manajemen Informatika dan Komputer ITMI, Medan, Indonesia

<sup>4</sup>Ilmu Komputer, Universitas Muhammadiyah OKU Timur, Ogan Komering Ulu Timur, Indonesia

Email: <sup>1\*</sup>hartatitammamahlubis@unprimdn.ac.id, <sup>2</sup>randy@itnb.ac.id, <sup>3</sup>muhammadnasrigea@gmail.com,

<sup>4</sup>ericariansafitri38@gmail.com

Email Penulis Korespondensi: hartatitammamahlubis@unprimdn.ac.id\*

Submitted: 04/06/2026; Accepted: 27/06/2026; Published: 30/06/2026

**Abstrak**– Depresi pada remaja merupakan salah satu masalah kesehatan mental yang dapat memengaruhi kesejahteraan psikologis, prestasi akademik, dan interaksi sosial. Identifikasi faktor-faktor yang berkontribusi terhadap risiko depresi diperlukan untuk mendukung upaya deteksi dini dan pencegahan yang lebih efektif. Penelitian ini bertujuan mengidentifikasi faktor dominan yang berkontribusi terhadap prediksi risiko depresi remaja menggunakan *Extreme Gradient Boosting* (XGBoost) dengan interpretasi *SHapley Additive exPlanations* (SHAP). Penelitian memanfaatkan dataset publik yang terdiri atas 1.200 sampel remaja dengan atribut karakteristik individu, kondisi psikologis, penggunaan media sosial, dan aspek gaya hidup. Untuk mengatasi ketidakseimbangan kelas digunakan pendekatan *scale\_pos\_weight*, sedangkan evaluasi model dilakukan menggunakan *Repeated Stratified Cross Validation*. Hasil pengujian menunjukkan rata-rata *F1-score* sebesar 96,11% dengan standar deviasi 5,66%, yang mengindikasikan performa klasifikasi yang baik dan konsisten. Analisis SHAP menunjukkan bahwa *sleep\_hours*, *stress\_level*, *daily\_social\_media\_hours*, dan *anxiety\_level* merupakan variabel yang memberikan kontribusi terbesar terhadap prediksi model. Temuan ini menunjukkan bahwa pada dataset yang digunakan, berkurangnya durasi tidur, tingginya tingkat stres dan kecemasan, serta penggunaan media sosial yang lebih intens berkontribusi terhadap peningkatan prediksi risiko depresi. Selain menghasilkan performa klasifikasi yang baik, pendekatan XGBoost dengan SHAP juga meningkatkan transparansi model dalam menjelaskan kontribusi masing-masing variabel terhadap hasil prediksi. Hasil penelitian ini bersifat prediktif berdasarkan dataset yang digunakan dan tidak dimaksudkan sebagai dasar diagnosis klinis depresi.

**Kata Kunci:** Depresi Remaja; XGBoost; *Explainable Artificial Intelligence*; SHAP; Klasifikasi.

**Abstract**– Adolescent depression is a significant mental health issue that can adversely affect psychological well-being, academic performance, and social interactions. Identifying the factors contributing to depression risk is essential to support more effective early detection and prevention strategies. This study aims to identify the dominant factors contributing to the prediction of adolescent depression risk using *Extreme Gradient Boosting* (XGBoost) interpreted through *SHapley Additive exPlanations* (SHAP). The study utilized a public dataset comprising 1,200 adolescent records, including individual characteristics, psychological conditions, social media usage, and lifestyle-related attributes. To address class imbalance, the *scale\_pos\_weight* approach was applied, while model performance was evaluated using *Repeated Stratified Cross-Validation*. The experimental results achieved a mean *F1-score* of 96.11% with a standard deviation of 5.66%, indicating robust and consistent classification performance. SHAP analysis revealed that *sleep\_hours*, *stress\_level*, *daily\_social\_media\_hours*, and *anxiety\_level* were the most influential features contributing to the model's predictions. The findings indicate that, within the dataset used, shorter sleep duration, higher levels of stress and anxiety, and more intensive social media use were associated with an increased predicted risk of depression. In addition to achieving strong predictive performance, the integration of XGBoost and SHAP improved model transparency by explaining the contribution of each feature to the prediction outcomes. The findings are predictive in nature, based on the dataset used, and should not be interpreted as a clinical diagnosis of depression.

**Keywords:** Adolescent Depression; XGBoost; *Explainable Artificial Intelligence*; SHAP; Classification

## 1. PENDAHULUAN

Salah satu gangguan kesehatan mental yang umum terjadi pada remaja adalah depresi, yang dapat memberikan pengaruh terhadap berbagai aspek kehidupan. Selain berdampak pada kesejahteraan psikologis, kondisi ini juga berhubungan dengan penurunan prestasi akademik, terganggunya hubungan sosial, dan perkembangan emosional pada tahap transisi menuju kedewasaan [1]. Yoo et al. melaporkan bahwa depresi menjadi salah satu penyebab utama disabilitas pada remaja dan sering kali berkembang sebelum memperoleh diagnosis klinis yang memadai [2]. Tren peningkatan kejadian depresi pada remaja yang terus terlihat dalam beberapa tahun terakhir menunjukkan bahwa masalah kesehatan mental telah menjadi perhatian utama dalam bidang kesehatan masyarakat [3].

Di Indonesia, kondisi kesehatan mental remaja juga menjadi perhatian penting. Temuan Indonesia National Adolescent Mental Health Survey (I-NAMHS) memperlihatkan bahwa masalah kesehatan mental masih banyak dijumpai pada kelompok remaja, sementara akses terhadap layanan profesional belum mampu menjangkau seluruh individu yang membutuhkan bantuan [4]. Kondisi tersebut menunjukkan pentingnya upaya deteksi dini

untuk mengurangi dampak jangka panjang yang dapat ditimbulkan oleh depresi. Do et al. mengidentifikasi pengalaman stres dan lingkungan sekolah sebagai faktor yang berkontribusi terhadap munculnya depresi pada remaja [5]. Chen et al. menemukan bahwa kualitas tidur, kondisi keluarga, serta latar belakang pendidikan merupakan variabel yang berkaitan erat dengan peningkatan kerentanan depresi pada remaja [6]. Yoo et al. menjelaskan bahwa faktor keluarga, termasuk riwayat depresi pada orang tua dan paparan pengalaman penuh tekanan sejak masa kanak-kanak, dapat meningkatkan kemungkinan munculnya depresi pada periode remaja [2].

Selain pengaruh faktor keluarga dan lingkungan sosial, berbagai aspek gaya hidup juga berkontribusi terhadap kondisi kesehatan mental remaja. Agyapong-Opoku et al. menemukan bahwa penggunaan media sosial dengan intensitas tinggi berkaitan dengan meningkatnya kecenderungan depresi dan kecemasan, terutama ketika aktivitas tersebut dilakukan secara pasif dan tanpa pengendalian yang memadai [7]. Temuan serupa dilaporkan oleh Syaqui dan Nahariani yang menunjukkan bahwa semakin sering remaja menggunakan media sosial, semakin besar pula kemungkinan munculnya gangguan psikologis, termasuk stres dan depresi [8]. Di sisi lain, pola tidur juga menjadi faktor yang berhubungan erat dengan kesehatan mental. Gao et al. mengemukakan bahwa hubungan antara durasi tidur dan depresi bersifat timbal balik, sehingga gangguan tidur tidak hanya dapat meningkatkan risiko depresi, tetapi juga dapat muncul sebagai dampak dari kondisi depresi itu sendiri [9]. Penelitian Owusu Mensah et al. menunjukkan bahwa masalah tidur merupakan kondisi yang umum dijumpai pada remaja dengan depresi dan dapat memperburuk gejala psikologis yang dialami [10]. Studi yang dilakukan oleh Kania et al. mengungkapkan adanya hubungan statistik yang signifikan antara kualitas tidur dan kesehatan mental remaja. Semakin rendah kualitas tidur yang dimiliki, semakin tinggi kecenderungan munculnya gejala stres, kecemasan, dan depresi [11].

Kompleksitas hubungan antar faktor risiko menyebabkan identifikasi dini depresi menjadi tantangan tersendiri apabila hanya menggunakan pendekatan statistik tradisional. Oleh karena itu, berbagai penelitian mulai memanfaatkan metode *machine learning* untuk membantu proses prediksi secara lebih efektif. Mardini et al. mengembangkan model berbasis XGBoost untuk mendeteksi depresi dan kecemasan pada remaja menggunakan data kesehatan dan faktor sosial dengan performa yang baik [12]. Xin et al. menunjukkan bahwa algoritma *machine learning* mampu mengenali faktor-faktor penting yang memengaruhi kondisi kesehatan mental pada anak dan remaja melalui analisis berbagai variabel secara bersamaan [13].

Kemajuan teknologi kecerdasan buatan membuka peluang yang lebih luas bagi penerapan metode *machine learning* untuk mendukung analisis dan prediksi pada bidang kesehatan mental. Do et al. menunjukkan bahwa *machine learning* mampu mengidentifikasi faktor-faktor risiko utama yang berkontribusi terhadap depresi remaja serta membantu proses deteksi dini melalui analisis multidimensi berbagai faktor psikologis dan lingkungan [5]. Sharma et al. melaporkan bahwa penerapan kecerdasan buatan pada kesehatan mental remaja telah digunakan untuk diagnosis, pemantauan gejala, hingga prediksi gangguan mental dengan hasil yang menjanjikan [14]. Selain itu, Negara et al. menunjukkan bahwa model *machine learning* berbasis data perilaku digital mampu digunakan sebagai sistem deteksi dini depresi secara efektif dan berpotensi mendukung layanan kesehatan mental yang lebih mudah diakses oleh masyarakat [15].

Salah satu algoritma yang banyak digunakan dalam penelitian kesehatan mental adalah *Extreme Gradient Boosting* (XGBoost). XGBoost dikenal mampu mengolah data dengan karakteristik kompleks, termasuk jumlah fitur yang besar, hubungan variabel yang tidak linier, dan distribusi kelas yang tidak seimbang [12]. Berdasarkan penelitian Chen et al., penggunaan algoritma XGBoost menunjukkan kemampuan prediktif yang lebih baik dalam mengidentifikasi risiko depresi pada remaja dibandingkan beberapa pendekatan *machine learning* alternatif [6]. Dalam penelitian yang dilakukan oleh Ahmed et al., algoritma XGBoost terbukti lebih efektif untuk klasifikasi depresi karena menghasilkan nilai akurasi yang melampaui model-model pembandingan yang digunakan [16].

Meskipun memiliki performa prediksi yang tinggi, model *machine learning* sering kali bersifat *black-box* sehingga sulit menjelaskan alasan di balik keputusan yang dihasilkan. Dalam konteks kesehatan mental, transparansi model menjadi aspek penting karena hasil prediksi dapat digunakan sebagai dasar dalam penyusunan strategi pencegahan dan intervensi. Mimikou et al. menegaskan bahwa pendekatan *Explainable Artificial Intelligence* (XAI) diperlukan untuk meningkatkan interpretabilitas model prediksi depresi [3]. Lamba et al. menunjukkan bahwa metode *SHapley Additive exPlanations* (SHAP) mampu mengidentifikasi faktor-faktor yang paling berpengaruh terhadap risiko depresi secara konsisten [17]. Menurut Molnar, SHAP merupakan salah satu metode interpretabilitas yang banyak digunakan karena mampu memberikan penjelasan kontribusi fitur baik pada tingkat global maupun lokal secara konsisten [18]. Meskipun demikian, sebagian besar penelitian terdahulu masih berfokus pada peningkatan performa prediksi tanpa memberikan interpretasi yang komprehensif mengenai kontribusi masing-masing faktor risiko. Selain itu, penelitian yang secara khusus mengombinasikan *Extreme Gradient Boosting* (XGBoost) dan SHAP untuk mengidentifikasi faktor dominan risiko depresi pada populasi remaja masih terbatas. Kondisi tersebut menunjukkan perlunya pengembangan model yang tidak hanya akurat dalam melakukan prediksi, tetapi juga mampu memberikan penjelasan yang transparan terhadap faktor-faktor yang memengaruhi risiko depresi.

Berdasarkan latar belakang tersebut, penelitian ini mengembangkan model prediksi depresi remaja menggunakan *Extreme Gradient Boosting* (XGBoost) yang dikombinasikan dengan *SHapley Additive exPlanations* (SHAP) untuk mengevaluasi performa klasifikasi sekaligus mengidentifikasi faktor-faktor yang memberikan kontribusi terbesar terhadap risiko depresi. Penelitian ini diharapkan dapat memberikan kontribusi

dalam pengembangan model prediksi depresi yang tidak hanya memiliki performa yang baik, tetapi juga mampu menjelaskan pengaruh masing-masing variabel secara transparan. Selain itu, hasil identifikasi faktor dominan yang diperoleh dapat menjadi sumber informasi bagi pendidik, orang tua, tenaga kesehatan, dan pembuat kebijakan dalam merancang strategi deteksi dini, pencegahan, serta penanganan kesehatan mental remaja yang lebih tepat sasaran.

## 2. METODOLOGI PENELITIAN

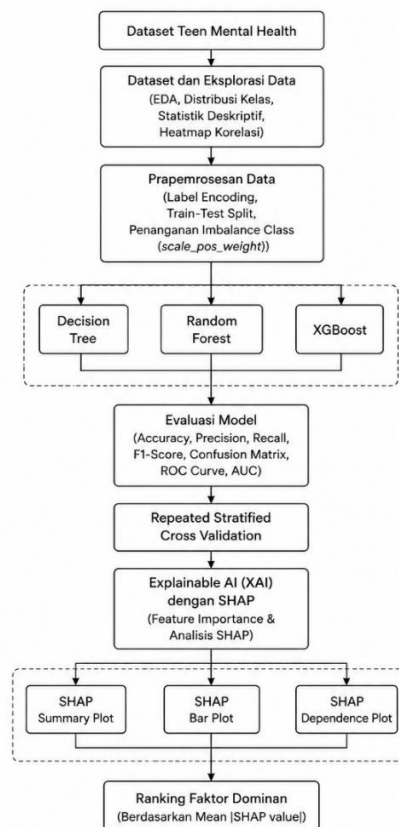
### 2.1 Tahapan Penelitian

Penelitian ini memanfaatkan Teenager Mental Health Dataset yang diperoleh dari platform Kaggle sebagai sumber data utama. Dataset tersebut tersedia secara publik pada tautan <https://www.kaggle.com/dataset/algozee/teenager-menthal-healy> dan dataset tersebut memuat berbagai atribut yang berkaitan dengan karakteristik individu, kebiasaan penggunaan media sosial, aktivitas fisik, kondisi psikologis, pola tidur, dan status depresi remaja. Dataset terdiri atas 1.200 sampel dengan distribusi kelas sebanyak 1.169 data tidak depresi dan 31 data depresi.

Tahapan penelitian diawali dengan eksplorasi data (*Exploratory Data Analysis* atau EDA) untuk memahami karakteristik dataset, distribusi kelas, statistik deskriptif, dan hubungan antarvariabel melalui analisis korelasi. Selanjutnya dilakukan prapemrosesan data yang meliputi transformasi variabel kategorikal menggunakan *Label Encoding*, pembagian data menjadi data pelatihan dan data pengujian dengan rasio 80:20 menggunakan *stratified sampling*, serta penanganan ketidakseimbangan kelas menggunakan parameter *scale\_pos\_weight* pada algoritma XGBoost.

Tahap berikutnya adalah pembangunan model klasifikasi menggunakan algoritma Decision Tree, Random Forest, dan XGBoost. Untuk memperoleh evaluasi yang lebih stabil dan representatif, setiap model divalidasi menggunakan metode *Repeated Stratified K-Fold Cross Validation*. Metode ini dipilih karena mampu mempertahankan proporsi kelas pada setiap fold sehingga sesuai untuk dataset dengan distribusi kelas yang tidak seimbang.

Kinerja model dievaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-Score*, *Area Under Curve* (AUC), *Confusion Matrix*, dan *Receiver Operating Characteristic* (ROC) Curve. Model dengan performa terbaik kemudian diinterpretasikan menggunakan pendekatan *Explainable Artificial Intelligence* (XAI) melalui metode *SHapley Additive exPlanations* (SHAP) untuk mengidentifikasi faktor-faktor dominan yang memengaruhi risiko depresi pada remaja. Alur lengkap penelitian ditunjukkan pada Gambar 1.



**Gambar 1.** Flowchart Alur Penelitian

### 2.1.1 Dataset Exploration

Penelitian ini menggunakan *Teenager Mental Health Dataset* yang diperoleh dari platform Kaggle [19]. Dataset tersebut merupakan dataset publik (*public dataset*) yang tersedia secara terbuka untuk keperluan penelitian dan pembelajaran. Dataset ini merupakan data sekunder yang berasal dari data survei mengenai kesehatan mental remaja serta memuat berbagai atribut yang berkaitan dengan karakteristik demografis, aktivitas sehari-hari, pola penggunaan media sosial, kualitas tidur, aktivitas fisik, interaksi sosial, tingkat stres, dan performa akademik. Variabel target pada penelitian ini adalah status depresi (*depression label*) yang digunakan untuk mengklasifikasikan remaja ke dalam kategori depresi dan tidak depresi.

Dataset terdiri atas 1.200 sampel dengan 13 variabel yang mencakup 12 variabel prediktor dan 1 variabel target. Untuk memberikan gambaran mengenai karakteristik data yang digunakan dalam penelitian, deskripsi setiap variabel disajikan pada Tabel 1.

**Tabel 1.** Karakteristik Variabel Dataset

Variabel	Tipe Data	Deskripsi
age	Integer	Usia responden
gender	Kategorikal	Jenis kelamin responden
academic_performance	Integer	Tingkat performa akademik
daily_social_media_hours	Float	Durasi penggunaan media sosial per hari
sleep_hours	Float	Durasi tidur per hari
stress_level	Integer	Tingkat stres responden
anxiety_level	Integer	Tingkat kecemasan responden
physical_activity	Integer	Tingkat aktivitas fisik
social_interaction_level	Integer	Tingkat interaksi sosial
addiction_level	Integer	Tingkat kecanduan media sosial
screen_time_before_sleep	Float	Durasi penggunaan layar sebelum tidur
platform_usage	Kategorikal	Platform media sosial yang dominan digunakan
depression	Integer	Variabel target (0 = Tidak Depresi, 1 = Depresi)

Tahap eksplorasi data dilakukan untuk memperoleh pemahaman awal mengenai karakteristik dataset sebelum proses pemodelan. Analisis yang dilakukan meliputi identifikasi jumlah atribut, tipe data, distribusi kelas, statistik deskriptif, analisis korelasi antarvariabel, pemeriksaan *missing values*, serta pengecekan konsistensi data. Hasil eksplorasi ini digunakan sebagai dasar dalam menentukan strategi prapemrosesan data yang sesuai sebelum dilakukan proses pembangunan model *machine learning*.

Berdasarkan hasil eksplorasi, distribusi kelas pada dataset menunjukkan kondisi *class imbalance*, yaitu jumlah sampel pada kelas tidak depresi jauh lebih besar dibandingkan kelas depresi. Distribusi kelas dataset disajikan pada Tabel 2.

**Tabel 2.** Distribusi Kelas Dataset

Kelas	Jumlah Sampel	Persentase
Tidak Depresi	1169	97,42%
Depresi	31	2,58%
Total	1200	100,00%

Berdasarkan Tabel 2, distribusi kelas yang tidak seimbang berpotensi menyebabkan model lebih cenderung memprediksi kelas mayoritas sehingga kemampuan dalam mengenali kelas minoritas menjadi menurun. Oleh karena itu, penelitian ini menerapkan parameter *scale\_pos\_weight* pada algoritma XGBoost untuk memberikan bobot yang lebih besar pada kelas minoritas selama proses pelatihan model sehingga pengaruh ketidakseimbangan kelas dapat dikurangi.

Karena menggunakan dataset publik yang telah dipublikasikan secara terbuka, penelitian ini tidak melibatkan pengumpulan data primer dari responden. Dataset yang digunakan tidak memuat informasi yang dapat mengidentifikasi identitas pribadi (*personally identifiable information*) sehingga privasi responden tetap terjaga. Dengan demikian, penelitian ini memanfaatkan data sekunder yang tersedia secara terbuka tanpa melakukan intervensi maupun pengambilan data langsung terhadap subjek penelitian. Selain itu, model yang dikembangkan dalam penelitian ini bertujuan untuk melakukan prediksi risiko depresi berdasarkan karakteristik data pada dataset yang digunakan, sehingga hasil yang diperoleh tidak dimaksudkan sebagai dasar diagnosis klinis depresi.

### 2.1.2 Prapemrosesan Data

Sebelum proses pemodelan dilakukan, data terlebih dahulu melalui tahap prapemrosesan untuk memastikan kesiapan data dalam proses pembelajaran mesin. Pada tahap ini, seluruh variabel kategorikal dikonversi ke bentuk numerik menggunakan teknik *Label Encoding*. Setelah itu, dataset dibagi menjadi data pelatihan dan data pengujian dengan perbandingan 80:20 menggunakan metode *stratified sampling* sehingga distribusi kelas pada kedua kelompok data tetap terjaga.

Berdasarkan hasil eksplorasi data, distribusi kelas pada dataset menunjukkan kondisi tidak seimbang (*class imbalance*), di mana jumlah sampel kategori tidak depresi jauh lebih besar dibandingkan kategori depresi. Ketidakseimbangan tersebut berpotensi menyebabkan model lebih cenderung memprediksi kelas mayoritas. Oleh karena itu, penelitian ini menerapkan parameter *scale\_pos\_weight* pada algoritma XGBoost untuk memberikan bobot yang lebih besar pada kelas minoritas.

Nilai *scale\_pos\_weight* dihitung berdasarkan rasio jumlah sampel kelas mayoritas terhadap jumlah sampel kelas minoritas sebagaimana ditunjukkan pada persamaan berikut:

$$scale\_pos\_weight = \frac{N_{mayoritas}}{N_{minoritas}} \tag{1}$$

Dimana:

$N_{mayoritas}$  : jumlah sampel kelas tidak depresi

$N_{minoritas}$  : jumlah sampel kelas depresi

Berdasarkan distribusi kelas pada dataset, jumlah sampel kategori tidak depresi adalah 1.169 data, sedangkan jumlah sampel kategori depresi adalah 31 data. Oleh karena itu, *scale\_pos\_weight* dihitung berdasarkan rasio kedua kelas tersebut dan diterapkan selama proses pelatihan model XGBoost untuk mengurangi bias terhadap kelas mayoritas serta meningkatkan kemampuan model dalam mengidentifikasi kasus depresi pada dataset yang tidak seimbang.

### 2.1.3 Pemodelan dan Evaluasi Model

Pada tahap pemodelan, penelitian ini menerapkan tiga algoritma klasifikasi, yaitu Decision Tree, Random Forest, dan XGBoost. Masing-masing algoritma dibangun menggunakan data pelatihan dan kemudian diuji menggunakan data pengujian untuk membandingkan kemampuan klasifikasinya dalam mendeteksi risiko depresi pada remaja. Hasil perbandingan tersebut digunakan untuk menentukan model dengan kinerja terbaik.

Untuk memastikan proses pemodelan dapat direplikasi, parameter utama yang digunakan pada algoritma XGBoost disajikan pada **Tabel 3**. Parameter tersebut digunakan selama proses pelatihan model untuk memperoleh performa klasifikasi yang optimal sekaligus mengurangi pengaruh ketidakseimbangan distribusi kelas.

**Tabel 3.** Parameter XGBoost

Parameter	Nilai
<b>n_estimators</b>	100
<b>max_depth</b>	3
<b>learning_rate</b>	0,05
<b>subsample</b>	0,7
<b>colsample_bytree</b>	0,7
<b>reg_alpha</b>	1
<b>reg_lambda</b>	2
<b>scale_pos_weight</b>	37,71
<b>random_state</b>	42
<b>eval_metric</b>	logloss

Nilai *scale\_pos\_weight* sebesar 37,71 diperoleh dari rasio jumlah sampel kelas mayoritas terhadap kelas minoritas. Parameter tersebut digunakan untuk mengurangi bias model terhadap kelas mayoritas sehingga kemampuan identifikasi kasus depresi dapat ditingkatkan. Sementara itu, parameter *eval\_metric* menggunakan *logloss* untuk mengukur kesalahan prediksi selama proses pelatihan model.

Kinerja setiap model dievaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-Score*, *Confusion Matrix*, ROC Curve, dan AUC. Perhitungan *Accuracy*, *Precision*, *Recall*, dan *F1-Score* didasarkan pada nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) yang diperoleh dari *confusion matrix*, sebagaimana ditunjukkan pada persamaan berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5}$$

Dimana:

- TP* : *True Positive*, yaitu jumlah data positif yang berhasil diprediksi dengan benar
- TN* : *True Negative*, yaitu jumlah data negatif yang berhasil diprediksi dengan benar
- FP* : *False Positive*, yaitu jumlah data negatif yang diprediksi sebagai positif
- FN* : *False Negative*, yaitu jumlah data positif yang diprediksi sebagai negatif

Selain itu, dilakukan perbandingan hasil pada data pelatihan dan data pengujian untuk menilai kemampuan generalisasi model serta mendeteksi kemungkinan terjadinya *overfitting*. Model dengan performa terbaik kemudian divalidasi menggunakan metode *Repeated Stratified K-Fold Cross Validation*. Pada penelitian ini digunakan 5 fold dan 10 pengulangan (*repeats*) dengan nilai *random\_state* sebesar 42. Pendekatan ini dipilih karena mampu mempertahankan proporsi kelas pada setiap *fold* sehingga sesuai digunakan pada dataset yang memiliki distribusi kelas tidak seimbang. Evaluasi validasi silang dilakukan menggunakan metrik *F1-Score*, dan nilai performa akhir diperoleh dari rata-rata seluruh hasil validasi silang sehingga menghasilkan estimasi kinerja model yang lebih stabil dan representatif.

### 2.1.4 Explainable Artificial Intelligence (XAI)

Model yang menghasilkan kinerja terbaik selanjutnya diinterpretasikan menggunakan pendekatan *Explainable Artificial Intelligence (XAI)* melalui metode *SHapley Additive exPlanations (SHAP)*. Metode SHAP digunakan untuk menjelaskan kontribusi setiap fitur terhadap hasil prediksi model berdasarkan nilai Shapley yang berasal dari teori permainan kooperatif [18].

Proses interpretasi dilakukan melalui beberapa visualisasi, yaitu *SHAP Summary Plot*, *SHAP Bar Plot*, dan *SHAP Dependence Plot*. *SHAP Summary Plot* digunakan untuk menunjukkan distribusi dan arah pengaruh setiap variabel terhadap hasil prediksi. *SHAP Bar Plot* digunakan untuk menampilkan tingkat kepentingan fitur berdasarkan nilai rata-rata absolut SHAP (*mean absolute SHAP value*), sedangkan *SHAP Dependence Plot* digunakan untuk menganalisis hubungan antara nilai suatu fitur dengan kontribusinya terhadap prediksi model.

Selanjutnya, faktor-faktor dominan risiko depresi ditentukan berdasarkan nilai rata-rata absolut SHAP sebagaimana ditunjukkan pada persamaan berikut:

$$Mean\ Absolute\ SHAP_j = \frac{1}{n} \sum_{i=1}^n |\phi_{ij}| \tag{6}$$

Dimana:

- Mean Absolute SHAP<sub>j</sub>* : nilai rata-rata absolut SHAP fitur ke-j
- $\phi_{ij}$  : nilai SHAP fitur ke-j pada observasi ke-i
- n* : jumlah observasi

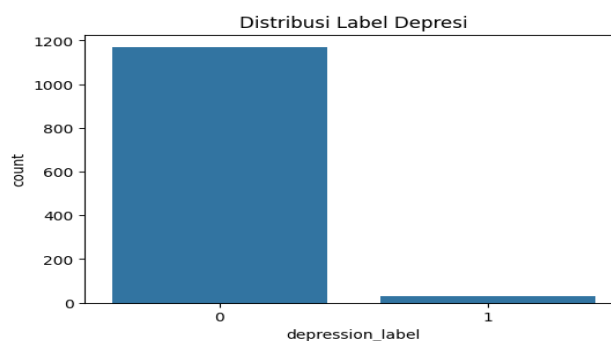
Variabel yang memiliki nilai *Mean Absolute SHAP Value* terbesar dianggap sebagai faktor yang memberikan kontribusi paling dominan terhadap prediksi risiko depresi pada remaja. Hasil analisis ini digunakan untuk mengidentifikasi faktor-faktor utama yang memengaruhi risiko depresi sekaligus meningkatkan interpretabilitas model sehingga hasil prediksi dapat dijelaskan secara lebih transparan.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Distribusi Data dan Analisis Awal

Analisis awal dilakukan guna mengeksplorasi karakteristik dataset sebelum digunakan dalam proses pembentukan model. Tahap ini mencakup pengkajian distribusi data target, analisis statistik deskriptif, serta identifikasi hubungan antarvariabel melalui korelasi. Informasi yang diperoleh dari tahap eksplorasi ini digunakan untuk memahami pola yang terdapat pada data sekaligus memberikan indikasi awal mengenai faktor-faktor yang berkaitan dengan depresi pada remaja.

Evaluasi distribusi kelas dilakukan guna memperoleh gambaran mengenai komposisi data pada setiap kategori target. Distribusi antara kelas depresi dan tidak depresi yang diperoleh dari dataset ditunjukkan pada Gambar 2.



Gambar 2. Distribusi Label Depresi

Berdasarkan distribusi data pada Gambar 2, mayoritas sampel berada pada kategori tidak depresi dengan jumlah 1.169 data, sedangkan kategori depresi hanya terdiri dari 31 data. Komposisi tersebut menunjukkan adanya ketimpangan distribusi kelas, di mana kelas tidak depresi mendominasi sebesar 97,42% dari keseluruhan data, sementara kelas depresi hanya menyumbang 2,58%.

Distribusi kelas yang tidak seimbang dapat memengaruhi proses pembelajaran model karena model cenderung lebih mudah mengenali pola pada kelas yang memiliki jumlah sampel lebih banyak. Untuk meminimalkan pengaruh kondisi tersebut, penelitian ini menerapkan *parameter scale\_pos\_weight* pada algoritma XGBoost. Pendekatan tersebut digunakan agar model dapat memberikan bobot yang lebih besar pada kelas depresi sehingga kemampuan dalam mengidentifikasi kasus depresi menjadi lebih optimal.

Selain analisis distribusi kelas, dilakukan pula analisis statistik deskriptif untuk menggambarkan karakteristik masing-masing kelompok berdasarkan status depresi. Hasil analisis tersebut disajikan pada Gambar 3.

```

---
Statistik Deskriptif Berdasarkan Label Depresi
age          daily_social_media_hours
  mean      std      mean      std
depression_label
0          15.924722  2.021582          4.478785  2.018025
1          16.064516  2.064550          6.719355  1.013384

sleep_hours  academic_performance
  mean      std      mean      std
depression_label
0          6.494183  1.432015          2.990248  0.577303
1          4.761290  0.566673          2.995484  0.564965

physical_activity  stress_level
  mean      std      mean      std
depression_label
0          1.016168  0.582658          5.365269  2.892040
1          0.951613  0.569720          8.483871  1.207503

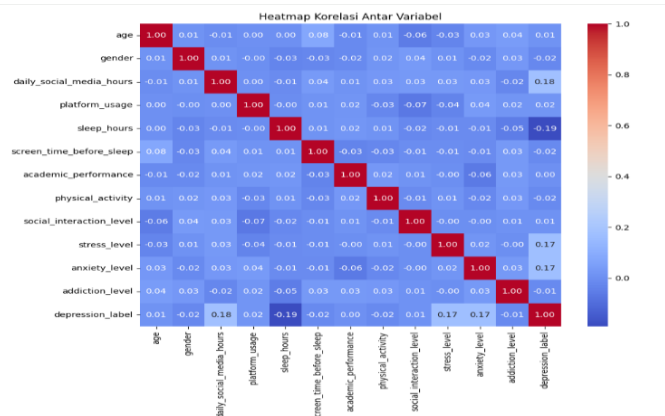
anxiety_level
  mean      std
depression_label
0          5.557742  2.850191
1          8.612903  1.054433
  
```

Gambar 3. Hasil Statistik Deskriptif

Berdasarkan Gambar 3, kelompok remaja yang mengalami depresi menunjukkan rata-rata durasi penggunaan media sosial yang lebih tinggi dibandingkan kelompok tidak depresi. Sebaliknya, rata-rata durasi tidur pada kelompok depresi cenderung lebih rendah. Perbedaan tersebut mengindikasikan adanya variasi pola aktivitas harian antara kedua kelompok.

Selain itu, variabel tingkat stres dan tingkat kecemasan juga memperlihatkan perbedaan yang cukup mencolok. Kelompok depresi memiliki rata-rata tingkat stres dan kecemasan yang lebih tinggi dibandingkan kelompok tidak depresi. Temuan ini menunjukkan bahwa kondisi psikologis dan gaya hidup sehari-hari berpotensi memiliki keterkaitan dengan munculnya risiko depresi pada remaja.

Untuk melengkapi analisis awal, dilakukan pengujian korelasi guna mengetahui hubungan antarvariabel yang terdapat dalam dataset. Hasil korelasi divisualisasikan menggunakan *heatmap* sebagaimana ditampilkan pada Gambar 4.



Gambar 4. Heatmap Korelasi Antar Variabel

Berdasarkan visualisasi korelasi pada Gambar 4, sebagian besar variabel menunjukkan tingkat hubungan yang rendah terhadap status depresi. Di antara seluruh variabel, *sleep\_hours* memiliki nilai korelasi negatif terbesar terhadap depresi, yaitu sebesar -0,19. Hasil tersebut menunjukkan bahwa penurunan durasi tidur cenderung berkaitan dengan meningkatnya risiko depresi pada remaja. Sebaliknya, *variabel daily\_social\_media\_hours*

memiliki korelasi positif sebesar 0,18, yang mengindikasikan bahwa semakin lama penggunaan media sosial, semakin tinggi kecenderungan munculnya depresi.

Variabel *stress\_level* dan *anxiety\_level* juga memperlihatkan hubungan positif dengan status depresi, masing-masing sebesar 0,17. Temuan ini menunjukkan bahwa peningkatan tingkat stres dan kecemasan berkaitan dengan peluang yang lebih besar untuk mengalami depresi. Meskipun demikian, nilai korelasi yang diperoleh masih tergolong rendah hingga sedang, sehingga hubungan yang teridentifikasi tidak menunjukkan keterkaitan yang sangat kuat.

Selain hubungan terhadap variabel target, hasil analisis juga menunjukkan bahwa tidak terdapat pasangan variabel independen yang memiliki korelasi sangat tinggi. Kondisi tersebut mengindikasikan bahwa risiko multikolinearitas pada dataset relatif kecil dan tidak diperkirakan akan memberikan pengaruh yang signifikan terhadap proses pemodelan. Secara keseluruhan, hasil analisis korelasi memberikan indikasi awal bahwa durasi tidur, intensitas penggunaan media sosial, tingkat stres, dan tingkat kecemasan merupakan faktor yang berpotensi berkontribusi terhadap risiko depresi pada remaja.

### 3.2 Perbandingan Kinerja Algoritma

Setelah melalui tahapan eksplorasi dan prapemrosesan data, dilakukan proses pemodelan menggunakan tiga algoritma klasifikasi, yaitu *Decision Tree*, *Random Forest*, dan *XGBoost*. Tujuan perbandingan ini adalah untuk menentukan algoritma yang paling efektif dalam mengidentifikasi risiko depresi pada remaja. Evaluasi kinerja model dilakukan menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Hasil perbandingan performa ketiga algoritma dapat dilihat pada Gambar 5.

	Model	Accuracy	Precision	Recall	F1 Score
0	Decision Tree	0.970833	0.461538	1.0	0.631579
1	Random Forest	0.987500	1.000000	0.5	0.666667
2	XGBoost	1.000000	1.000000	1.0	1.000000

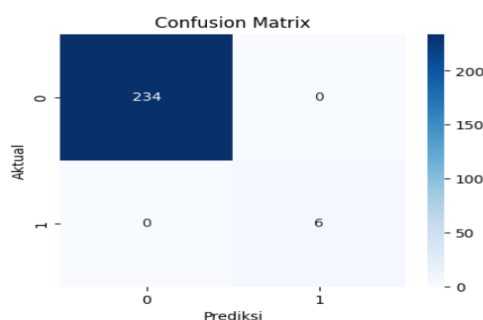
**Gambar 5.** Hasil Perbandingan Kinerja Algoritma

Hasil perbandingan pada Gambar 5, menunjukkan bahwa *XGBoost* merupakan algoritma dengan kinerja terbaik pada penelitian ini. Seluruh metrik evaluasi, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score*, mencapai nilai 1,00 pada data pengujian. Hasil tersebut menunjukkan kemampuan model dalam mengklasifikasikan seluruh data uji dengan benar. Meskipun demikian, hasil yang sangat tinggi perlu diinterpretasikan secara hati-hati karena dataset yang digunakan memiliki distribusi kelas yang tidak seimbang, yaitu 1.169 data tidak depresi dan 31 data depresi. Pada kondisi seperti ini, performa yang sangat tinggi dapat dipengaruhi oleh karakteristik data maupun ukuran sampel kelas minoritas yang relatif kecil. Oleh karena itu, diperlukan evaluasi tambahan melalui validasi silang untuk memperoleh gambaran performa model yang lebih representatif.

*Decision Tree* mampu mencapai *Recall* sebesar 1,00, namun nilai *Precision* yang relatif rendah sebesar 0,46 menunjukkan adanya kecenderungan menghasilkan prediksi positif yang berlebihan. Sebaliknya, *Random Forest* memiliki *Precision* sebesar 1,00, tetapi *Recall* sebesar 0,50 menunjukkan bahwa sebagian kasus depresi masih belum berhasil terdeteksi. Dibandingkan kedua algoritma tersebut, *XGBoost* menunjukkan keseimbangan performa yang lebih baik dalam mengidentifikasi risiko depresi pada remaja.

### 3.3 Evaluasi Model XGBoost

Setelah terpilih sebagai model dengan performa terbaik, *XGBoost* dievaluasi lebih lanjut untuk menilai kemampuan klasifikasi dan konsistensi kinerjanya. Salah satu metode yang digunakan adalah *Confusion Matrix* yang memberikan informasi mengenai jumlah prediksi yang sesuai maupun tidak sesuai dengan kelas sebenarnya. Hasil *Confusion Matrix* ditunjukkan pada Gambar 6.



**Gambar 6.** Confusion Matrix

Hasil confusion matrix yang ditampilkan pada Gambar 6, memperlihatkan bahwa seluruh data uji berhasil dipetakan ke kelas yang sesuai. Sebanyak 234 sampel tidak depresi diprediksi dengan benar sebagai kelas tidak depresi, sedangkan seluruh 6 sampel depresi juga berhasil dikenali sebagai kelas depresi. Tidak adanya kesalahan klasifikasi yang ditunjukkan oleh nilai False Positive dan False Negative sebesar nol mengindikasikan bahwa model memiliki performa yang sangat baik dalam mengidentifikasi risiko depresi pada data pengujian.

Untuk memperoleh evaluasi yang lebih komprehensif, dilakukan validasi menggunakan metode Repeated Stratified Cross Validation. Pendekatan ini digunakan untuk mengukur kestabilan model melalui beberapa skenario pembagian data yang berbeda. Hasil validasi ditampilkan pada Gambar 7.

Mean F1 : 0.9611421911421911

Std : 0.05658894525158025

**Gambar 7.** Hasil *Repeated Stratified Cross Validation*

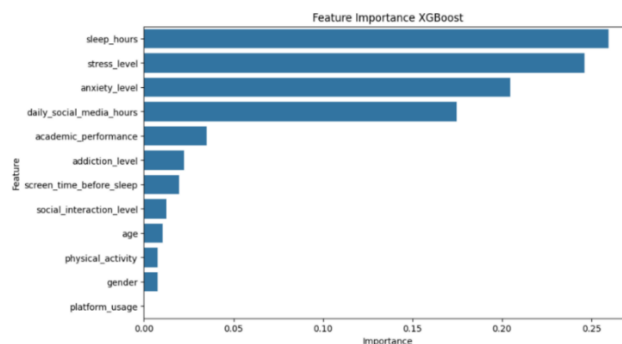
Gambar 7 menunjukkan bahwa, hasil validasi silang model XGBoost memperoleh nilai rata-rata *F1-Score* sebesar 96,11% dengan standar deviasi 5,66%. Nilai tersebut sedikit lebih rendah dibandingkan hasil evaluasi pada data pengujian yang mencapai nilai sempurna. Perbedaan ini menunjukkan bahwa performa model dapat bervariasi pada skenario pembagian data yang berbeda. Dengan demikian, hasil validasi silang memberikan gambaran yang lebih representatif mengenai kemampuan generalisasi model dibandingkan hanya mengandalkan satu pembagian data pengujian. Meskipun demikian, nilai *F1-Score* yang tetap tinggi menunjukkan bahwa model memiliki kemampuan klasifikasi yang baik dan stabil dalam mendeteksi risiko depresi pada remaja.

### 3.4 Analisis Explainable XGBoost

Setelah proses evaluasi menunjukkan bahwa XGBoost merupakan model dengan kinerja terbaik, dilakukan analisis lebih lanjut menggunakan metode *SHapley Additive exPlanations* (SHAP) dalam kerangka *Explainable Artificial Intelligence* (XAI). Pendekatan ini memungkinkan identifikasi kontribusi setiap fitur terhadap keputusan yang dihasilkan model, sehingga variabel-variabel yang paling berpengaruh terhadap risiko depresi dapat diketahui. Selain meningkatkan pemahaman terhadap mekanisme prediksi model, analisis SHAP juga membantu menghasilkan interpretasi yang lebih transparan dan informatif.

#### 3.4.1 Feature Importance

Analisis *Feature Importance* dilakukan untuk mengetahui tingkat kontribusi setiap variabel dalam proses klasifikasi depresi menggunakan model XGBoost. Hasil *Feature Importance* ditunjukkan pada Gambar 8.



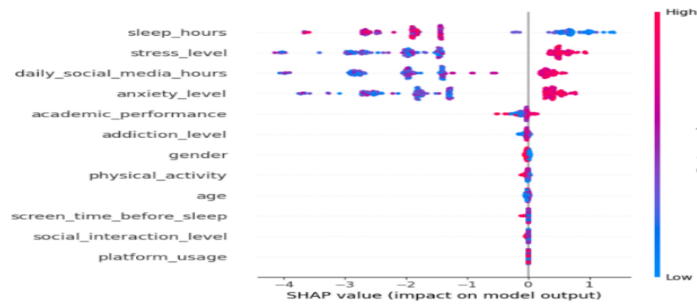
**Gambar 8.** Hasil *Feature Importance*

Berdasarkan Gambar 8, variabel *sleep\_hours* memiliki nilai importance tertinggi, diikuti oleh *stress\_level*, *anxiety\_level*, dan *daily\_social\_media\_hours*. Hasil ini menunjukkan bahwa durasi tidur, tingkat stres, tingkat kecemasan, dan penggunaan media sosial merupakan variabel yang paling sering digunakan model dalam membentuk keputusan klasifikasi. Sementara itu, variabel seperti *platform\_usage*, *social\_interaction\_level*, dan *screen\_time\_before\_sleep* memiliki kontribusi yang relatif rendah terhadap prediksi depresi.

Meskipun *Feature Importance* mampu menunjukkan tingkat kepentingan fitur, metode ini belum dapat menjelaskan arah pengaruh masing-masing variabel terhadap hasil prediksi. Oleh karena itu, dilakukan analisis lanjutan menggunakan SHAP.

#### 3.4.2 SHAP Summary Plot

SHAP Summary Plot digunakan untuk memberikan gambaran global mengenai kontribusi setiap variabel terhadap prediksi model sekaligus menunjukkan arah pengaruhnya. Hasil visualisasi SHAP Summary Plot ditunjukkan pada **Gambar 9**.

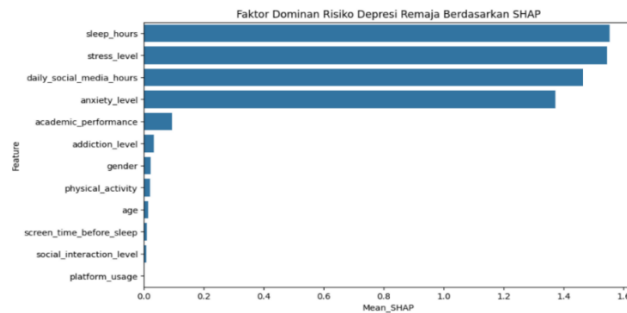


**Gambar 9.** Summary Plot

Berdasarkan Gambar 9, variabel *sleep\_hours*, *stress\_level*, *daily\_social\_media\_hours*, dan *anxiety\_level* merupakan fitur yang paling berpengaruh terhadap prediksi depresi. Nilai *sleep\_hours* yang rendah cenderung meningkatkan risiko depresi, sedangkan durasi tidur yang lebih tinggi berkontribusi menurunkan risiko. Sebaliknya, nilai yang tinggi pada *stress\_level*, *anxiety\_level*, dan *daily\_social\_media\_hours* cenderung meningkatkan probabilitas depresi. Hasil ini menunjukkan bahwa kurang tidur, tingginya stres, kecemasan, dan penggunaan media sosial merupakan faktor utama yang berkontribusi terhadap risiko depresi pada remaja.

### 3.4.3 Faktor Dominan Risiko Depresi Remaja

Untuk memudahkan interpretasi hasil, dilakukan visualisasi faktor dominan berdasarkan nilai rata-rata SHAP sebagaimana ditunjukkan pada Gambar 10.



**Gambar 10.** Faktor Dominan Risiko Depresi Remaja Berdasarkan SHAP

Berdasarkan Gambar 10, faktor yang paling dominan memengaruhi risiko depresi remaja adalah *sleep\_hours*, diikuti oleh *stress\_level*, *daily\_social\_media\_hours*, *anxiety\_level*, dan *academic\_performance*. Hasil ini menunjukkan bahwa faktor gaya hidup dan kondisi psikologis memiliki pengaruh lebih besar terhadap risiko depresi dibandingkan faktor demografis. Durasi tidur yang rendah, tingkat stres dan kecemasan yang tinggi, serta penggunaan media sosial yang berlebihan menjadi indikator utama dalam deteksi dini depresi pada remaja.

### 3.5 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma XGBoost mampu memberikan performa klasifikasi yang lebih baik dibandingkan Decision Tree dan Random Forest. Seluruh metrik evaluasi pada data pengujian menghasilkan nilai yang sangat tinggi, bahkan mencapai nilai sempurna sebesar 1,00. Meskipun demikian, hasil tersebut perlu diinterpretasikan secara hati-hati karena dataset yang digunakan memiliki distribusi kelas yang tidak seimbang, yaitu 1.169 data tidak depresi dan 31 data depresi. Kondisi class imbalance tersebut berpotensi menyebabkan nilai evaluasi pada data pengujian tampak sangat tinggi, terutama karena jumlah sampel pada kelas minoritas relatif sedikit. Oleh karena itu, penelitian ini tidak hanya mengandalkan evaluasi pada data pengujian, tetapi juga menerapkan *Repeated Stratified K-Fold Cross Validation* untuk memperoleh estimasi performa model yang lebih representatif dan mengurangi kemungkinan bias akibat pembagian data tertentu.

Hasil validasi silang menunjukkan bahwa model XGBoost memperoleh rata-rata *F1-Score* sebesar 96,11% dengan standar deviasi 5,66%. Nilai tersebut sedikit lebih rendah dibandingkan hasil evaluasi pada data pengujian, namun tetap menunjukkan performa klasifikasi yang tinggi dan relatif stabil pada berbagai skenario pembagian data. Temuan ini mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik terhadap

dataset yang digunakan serta tidak hanya bergantung pada satu pembagian data tertentu. Selain itu, penerapan parameter *scale\_pos\_weight* membantu mengurangi kecenderungan model untuk lebih memprioritaskan kelas mayoritas dengan memberikan bobot yang lebih besar pada kelas minoritas selama proses pelatihan.

Analisis interpretabilitas menggunakan SHAP menunjukkan bahwa *sleep\_hours*, *stress\_level*, *daily\_social\_media\_hours*, dan *anxiety\_level* merupakan variabel yang memberikan kontribusi terbesar terhadap hasil prediksi model. Berdasarkan interpretasi SHAP, berkurangnya durasi tidur, meningkatnya tingkat stres dan kecemasan, serta intensitas penggunaan media sosial yang lebih tinggi memiliki kontribusi yang lebih besar terhadap prediksi risiko depresi pada dataset yang digunakan. Temuan tersebut menunjukkan bahwa faktor psikologis dan pola hidup menjadi variabel yang paling dominan dalam proses prediksi model. Namun demikian, interpretasi SHAP bersifat menjelaskan kontribusi setiap variabel terhadap hasil prediksi model dan tidak dapat diartikan sebagai hubungan sebab-akibat maupun diagnosis klinis depresi.

Hasil penelitian ini memiliki kesesuaian dengan studi yang dilakukan oleh Junianto dan Nurkhodijah yang menerapkan pendekatan *Explainable Ensemble Learning* untuk klasifikasi risiko depresi berbasis *machine learning* dan SHAP. Penelitian tersebut melaporkan bahwa masalah tidur, tingginya penggunaan media sosial, durasi screen time yang berlebihan, serta indikator psikologis lainnya merupakan faktor yang berkontribusi terhadap peningkatan risiko depresi. Melalui analisis SHAP, variabel *sleep issues*, *social media hours*, *total screen time*, dan *scroll rate* juga teridentifikasi sebagai prediktor penting pada kelompok dengan tingkat risiko depresi sedang hingga tinggi [20]. Temuan penelitian ini juga sejalan dengan Chen et al. [6] yang menunjukkan bahwa kualitas tidur dan kondisi psikologis merupakan faktor penting dalam identifikasi risiko depresi pada remaja. Selain itu, Mardini et al. [12] melaporkan bahwa algoritma XGBoost mampu menghasilkan performa yang baik dalam deteksi depresi dan kecemasan berdasarkan kombinasi faktor sosial dan kesehatan. Kesamaan hasil tersebut menunjukkan bahwa durasi tidur, tingkat stres, kecemasan, dan aktivitas digital merupakan faktor yang secara konsisten muncul sebagai prediktor penting pada berbagai penelitian terkait kesehatan mental remaja.

Secara keseluruhan, hasil penelitian menunjukkan bahwa faktor psikologis dan pola hidup memberikan kontribusi yang lebih besar terhadap prediksi model dibandingkan karakteristik demografis, seperti usia dan jenis kelamin. Temuan ini mengindikasikan bahwa kualitas tidur, tingkat stres, kecemasan, serta penggunaan media sosial merupakan variabel yang layak diperhatikan dalam pengembangan sistem pendukung deteksi dini berbasis *machine learning*. Selain menghasilkan performa klasifikasi yang baik, pendekatan *Explainable XGBoost* memberikan transparansi dalam menjelaskan kontribusi setiap variabel terhadap hasil prediksi sehingga dapat membantu peneliti maupun praktisi memahami dasar pengambilan keputusan model. Meskipun demikian, hasil penelitian ini masih terbatas pada karakteristik dataset yang digunakan dan memerlukan validasi lebih lanjut menggunakan dataset lain maupun data klinis untuk menguji kemampuan generalisasi model pada populasi yang lebih luas.

#### 4. KESIMPULAN

Penelitian ini menunjukkan bahwa kombinasi algoritma *Extreme Gradient Boosting* (XGBoost) dengan pendekatan *SHapley Additive exPlanations* (SHAP) dapat digunakan untuk mengidentifikasi faktor-faktor yang memberikan kontribusi terhadap prediksi risiko depresi pada remaja berdasarkan dataset yang digunakan. Berdasarkan hasil evaluasi model, XGBoost memperoleh nilai rata-rata *F1-Score* sebesar 96,11% dengan standar deviasi 5,66% pada proses *Repeated Stratified K-Fold Cross Validation*. Hasil tersebut menunjukkan bahwa model memiliki performa klasifikasi yang baik dan relatif konsisten pada berbagai skenario pembagian data. Hasil interpretasi menggunakan SHAP menunjukkan bahwa variabel *sleep\_hours*, *stress\_level*, *daily\_social\_media\_hours*, dan *anxiety\_level* merupakan variabel yang memberikan kontribusi terbesar terhadap hasil prediksi model. Temuan ini mengindikasikan bahwa pada dataset yang digunakan, durasi tidur yang lebih rendah, tingkat stres dan kecemasan yang lebih tinggi, serta penggunaan media sosial yang lebih intens berkontribusi terhadap peningkatan prediksi risiko depresi. Namun demikian, interpretasi SHAP menjelaskan kontribusi setiap variabel terhadap hasil prediksi model sehingga tidak dapat diartikan sebagai hubungan sebab-akibat (*causal relationship*) maupun digunakan sebagai dasar diagnosis klinis depresi. Meskipun menghasilkan performa klasifikasi yang baik, penelitian ini masih memiliki beberapa keterbatasan. Dataset yang digunakan memiliki distribusi kelas yang tidak seimbang dengan proporsi kelas depresi yang jauh lebih sedikit dibandingkan kelas tidak depresi sehingga berpotensi memengaruhi proses pembelajaran model. Selain itu, penggunaan dataset publik sebagai data sekunder membatasi kemampuan penelitian dalam merepresentasikan kondisi populasi remaja secara lebih luas. Oleh karena itu, hasil penelitian ini masih memerlukan validasi lebih lanjut menggunakan dataset lain maupun data klinis yang lebih beragam untuk menguji kemampuan generalisasi model. Penelitian selanjutnya disarankan menggunakan dataset dengan jumlah sampel yang lebih besar, distribusi kelas yang lebih seimbang, serta karakteristik responden yang lebih beragam agar model yang dihasilkan menjadi lebih robust. Selain itu, penerapan dan perbandingan berbagai metode *Explainable Artificial Intelligence* (XAI), seperti SHAP, LIME,

maupun metode interpretabilitas lainnya, dapat dilakukan untuk memperkaya interpretasi model serta mengevaluasi konsistensi faktor-faktor yang berkontribusi terhadap prediksi risiko depresi pada remaja.

## REFERENCES

- [1] A. Maerlender and P. G. Bray, *Adolescent Mental Health in the Twenty-First Century.pdf*, 1st Editio. Cham, Switzerland: Springer Nature, 2024. doi: 10.1007/978-3-031-68294-0\_2.
- [2] A. Yoo *et al.*, "Prediction of adolescent depression from prenatal and childhood data from ALSPAC using machine learning," *Sci. Rep.*, pp. 1–13, 2024, doi: 10.1038/s41598-024-72158-9.
- [3] C. Mimikou *et al.*, "Explainable Machine Learning in the Prediction of Depression," *Diagnostics*, vol. 15, no. 11, pp. 1–18, 2025, doi: 10.3390/diagnostics15111412.
- [4] Amirah Ellyza Wahdi; Althaf Setyawan; Yufan Astrini Putri; Siswanto Agus Wilopo; Holly E. Erskine; Krystina Wallis; Cartiah McGrath; Sarah J. Blondell; Harvey A. Whiteford; James G. Scott; Robert Blum; Shoshanna Fine; Mengmeng Li; Astha Ramaiya, "Indonesia – National Adolescent Mental Health Survey (I-NAMHS) Report," Center for Reproductive Health, Yogyakarta, Indonesia, 2022.
- [5] H. Do *et al.*, "Apply Machine Learning to Predict Risk for Adolescent Depression in a Cohort of Kenyan Adolescents," *Healthcare*, vol. 13, no. 20, pp. 1–19, 2025, doi: 10.3390/healthcare13202620.
- [6] X. Chen *et al.*, "BMC Med Inform Decis Mak Article in Press Predicting adolescent depression : an interpretable machine learning model IN AR IN," *BMC Med. Inform. Decis. Mak.*, 2026, doi: 10.1186/s12911-026-03482-5.
- [7] N. Agyapong-opoku, F. Agyapong-opoku, and A. J. Greenshaw, "Effects of Social Media Use on Youth and Adolescent Mental Health : A Scoping Review of Reviews," *Behav. Sci. (Basel).*, vol. 15, no. 5, pp. 1–22, 2025, doi: 10.3390/bs15050574.
- [8] M. Syahrizzidan, A. Syauqi, and P. Nahariani, "The Impact of Social Media Use on Adolescent Mental Health," *J. Ilm. Keperawatan*, vol. 20, pp. 40–47, 2025, doi: 10.30643/jiksh.v20i2.427.
- [9] M. Gao and C. Lee, "Sleep duration and depression among adolescents : Mediation effect of collective integration," no. November, pp. 1–14, 2022, doi: 10.3389/fpsyg.2022.1015089.
- [10] S. O. Mensah *et al.*, "Evaluation and treatment of sleep disorders in adolescents with depression ; a scoping review," *BMC Psychiatry*, vol. 25, no. 2025, 2026, doi: 10.1186/s12888-025-07652-z.
- [11] A. Putri and A. Ispriantari, "Sleep Quality And Mental Health Among Junior High School Students," *Prof. Heal. J.*, vol. 8, no. 1, pp. 125–132, 2026, doi: 10.54832/phj.v8i1.1347.
- [12] M. T. Mardini, G. E. Khalil, C. Bai, A. M. Divakaran, and J. M. Ray, "Identifying Adolescent Depression and Anxiety Through Real-World Data and Social Determinants of Health : Machine Learning Model Development and Validation," vol. 12, pp. 1–11, 2025, doi: 10.2196/66665.
- [13] Y. Xin *et al.*, "Machine learning-based analysis and prediction of factors influencing mental health among children and adolescents in Jiangsu Province," *Child Adolesc. Psychiatry Ment. Health*, vol. 5, pp. 1–8, 2025, doi: 10.1186/s13034-025-00959-5.
- [14] G. Sharma, M. J. Yaffe, P. Ghadiri, R. Gandhi, and L. Pinkham, "Use of Artificial Intelligence in Adolescents ' Mental Health Care : Systematic Scoping Review of Current Applications and Future Directions," vol. 12, pp. 1–14, doi: 10.2196/70438.
- [15] J. Ilmiah *et al.*, "EARLY DETECTION OF ACADEMIC DEPRESSION USING SMARTPHONE-," *J. Ilm. Ilmu Terap. Univ. Jambi*, vol. 9, no. 3, pp. 1104–1116, 2025, doi: 10.22437/jiituj.v9i3.46375.
- [16] I. Ahmed, A. Brahmacharimayum, R. H. Ali, T. A. Khan, M. O. Ahmad, and M. O. Ahmad, "Explainable AI for Depression Detection and Severity Classification From Activity Data : Development and Evaluation Study of an Interpretable Framework," *JMIR Ment. Heal.*, vol. 12, no. 1, 2025, doi: 10.2196/72038.
- [17] K. Lamba, S. Rani, and M. Shabaz, "Explainable machine learning for mental health prediction from social media behavior : a nested cross-validation study with SHAP and LIME interpretability," *Discov. Ment. Heal.*, vol. 6, 2026, doi: 10.1007/s44192-026-00373-z.
- [18] C. Molnar, *Interpretable Machine Learning A Guide for Making Black Box Models Explainable*. Munich, Germany: Leanpub, 2022.
- [19] M. Shahzad, "Social Media Impact on Teen Mental Health," Kaggle. Accessed: Jun. 02, 2026. [Online]. Available: <https://www.kaggle.com/datasets/algozee/teenager-menthal-healy>
- [20] E. Junianto and S. Nurkhodijah, "Explainable Ensemble Learning for Depression Risk Classification Using Multidomain Behavioral Features," *J. Tek. Inform.*, vol. 7, no. 2, pp. 778–792, 2026, doi: 10.52436/1.jutif.2026.7.2.5009.