

Perbandingan Kinerja Algoritma *Machine Learning Logistic Regression* dan *XGBoost* dalam Memprediksi Gangguan Tidur dengan Analisis *SHAP*

Reza Fitriansyah¹, Ahmad Fauzan Baihaqi²,

Jurusan Teknologi Informasi, ITB Ahmad Dahlan, Jl. Ir H. Juanda No.77, Cireundeu, Kec. Ciputat Tim., Kota Tangerang Selatan, Banten 15419¹

Jurusan Sejarah, Universitas Diponegoro, Jl. Prof. Soedarto, Tembalang, Kec. Tembalang, Kota Semarang, Jawa Tengah 50275²

rezajurnal91@gmail.com¹, ahmadfauzanbaihaqi25@lecturer.undip.ac.id²

Abstrak Gangguan tidur merupakan masalah kesehatan masyarakat yang berdampak pada produktivitas, kesehatan fisik, dan kesejahteraan mental. Penelitian ini bertujuan untuk memprediksi faktor-faktor yang memengaruhi gangguan tidur dengan membandingkan kinerja algoritma *Logistic Regression* dan *Extreme Gradient Boosting (XGBoost)*. Data diperoleh dari survei *National Sleep Foundation (NSF)* dan melalui tahap pra-pemrosesan sebelum digunakan untuk pemodelan. Evaluasi dilakukan menggunakan metrik akurasi, presisi, *Recall*, dan *F1-score*, serta interpretabilitas model ditingkatkan dengan metode *SHAP (SHAPley Additive exPlanations)*. Hasil penelitian menunjukkan bahwa *XGBoost* unggul dengan akurasi, presisi, *Recall*, dan *F1-score* sebesar 100%, sedangkan *Logistic Regression* hanya mencapai 70%. Analisis *SHAP* mengungkapkan bahwa pola tidur ($\pm 1,9$), usia ($\pm 1,5$), konsumsi kafein ($\pm 1,3$), dan aktivitas fisik ($\pm 1,25$) merupakan faktor dominan, sementara kondisi kesehatan umum berpengaruh sedang ($\pm 1,1$), dan jenis kelamin ($\pm 0,15$) serta kesehatan subjektif ($\pm 0,05$) berkontribusi minimal. Temuan ini menegaskan bahwa gaya hidup dan kebiasaan harian menjadi penentu utama kualitas tidur. Ke depan, penelitian dapat diperluas dengan dataset yang lebih besar dan integrasi data fisiologis *real-time* dari *wearable device* untuk menghasilkan sistem prediksi yang lebih generalis dan aplikatif dalam memberikan rekomendasi personal terkait pola tidur sehat.

Kata Kunci: Gangguan Tidur, *Machine Learning*, *XGBoost*, *Logistic Regression*, *SHAP*

I. PENDAHULUAN

Gangguan tidur kini menjadi isu kesehatan masyarakat yang semakin sering diperbincangkan karena berdampak luas pada kesejahteraan individu, produktivitas, serta kondisi fisik maupun mental. Menurut *National Sleep Foundation (NSF)*, kualitas tidur dipengaruhi oleh berbagai aspek, seperti keadaan psikologis, kebiasaan hidup, aktivitas harian, hingga status kesehatan seseorang. Gangguan tidur tidak hanya menurunkan fokus dan performa kerja, tetapi juga dapat meningkatkan kerentanan terhadap penyakit kronis, misalnya hipertensi, diabetes, obesitas, serta gangguan kejiwaan seperti kecemasan dan depresi [1].

Berbagai studi menunjukkan bahwa kualitas tidur erat kaitannya dengan sejumlah faktor, antara lain tingkat stres, keteraturan jadwal tidur, konsumsi kafein, aktivitas fisik, status kesehatan, usia, dan jenis kelamin. Mengetahui faktor-faktor tersebut menjadi penting karena dapat memberikan pemahaman menyeluruh mengenai penyebab gangguan tidur sekaligus menjadi landasan dalam menyusun strategi pencegahan maupun intervensi yang lebih efektif [2].

Penelitian terbaru terkait gangguan tidur juga semakin banyak menggunakan pendekatan berbasis data melalui analisis prediktif. Penggunaan model prediksi dengan *Machine Learning* dinilai lebih akurat dalam memetakan faktor risiko dibandingkan metode tradisional. Hasil sejumlah penelitian

menunjukkan bahwa stres psikologis, tingginya konsumsi kafein, serta rendahnya tingkat aktivitas fisik merupakan indikator utama munculnya insomnia dan gangguan tidur lainnya[3].

National Sleep Foundation menegaskan bahwa keteraturan pola tidur, penerapan gaya hidup sehat, dan kondisi fisik yang terjaga merupakan kunci utama untuk memperoleh kualitas tidur yang baik. Sejalan dengan itu, pemanfaatan analisis multivariat dalam penelitian memungkinkan peneliti mengidentifikasi hubungan kompleks antarvariabel yang seringkali sulit ditangkap dengan metode konvensional.

II. TINJAUAN PUSTAKA

1. Aziz et al. (2024)

Melakukan scoping review pada 46 studi *wearable* berbasis *AI*. Algoritma terbanyak: *CNN* (36,9%), *Random Forest* (30,4%), dan *SVM* (26,1%). Fokus utama deteksi sleep apnea dengan data respirasi dan detak jantung[4].

2. Khanmohmmadi et al. (2025)

Mengembangkan *multi-task learning* berbasis *CNN* dengan optimasi *genetic algorithm* dan *Q-learning* untuk *EEG*, mencapai akurasi 98% dalam mendeteksi *partial sleep deprivation* [5].

3. Septiadi & Prasetyo (2025)

Menggunakan *K-Means clustering* pada 374 responden, menghasilkan 10 *cluster* yang menunjukkan hubungan erat aktivitas fisik, stres, dan *BMI* dengan kualitas tidur[6].

4. Guyot et al. (2025)

Menerapkan algoritma berbasis *ECG* jangka panjang untuk *sleep apnea*, mencapai sensitivitas 93,3% dan spesifisitas 66,7%, menunjukkan *ECG* berpotensi jadi alternatif *PSG*[7].

5. Monowar et al. (2025)

Mengusulkan *multi-layered ensemble learning* dengan *SMOTE*, meraih akurasi 96,88%, lebih tinggi dari model tunggal dan efektif menangani data tidak seimbang[8].

6. Iqbal (2024)

Menggunakan algoritma *Decision Tree* untuk klasifikasi gangguan tidur dengan akurasi 85%. Keunggulan terletak pada interpretasi aturan yang sederhana, namun kelemahan pada risiko *overfitting*. Disarankan

eksplorasi algoritma ensemble untuk hasil lebih baik [9].

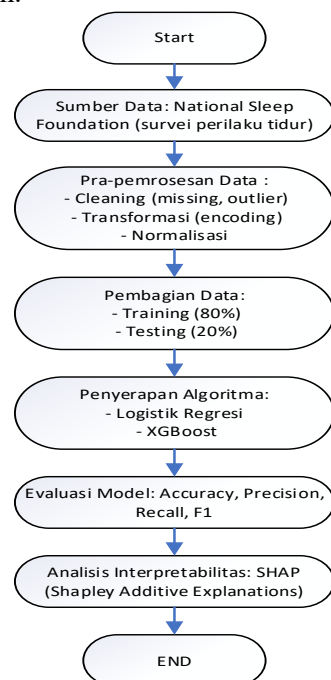
7. Khasanah et al. (2025)

Membandingkan *Random Forest* dan *KNN* dalam klasifikasi gangguan tidur. *Random Forest* lebih unggul dengan akurasi 89,69% dan *recall* 96,08% pada kategori *None*, sementara *KNN* hanya mencapai 87,02%. Kendala utama ada pada deteksi *Sleep Apnea* yang rendah akibat data *imbalance*, sehingga diperlukan metode *balancing* dan algoritma alternatif [10].

III. ANALISA DAN PERANCANGAN SISTEM

Studi ini mengadopsi pendekatan kuantitatif dengan memanfaatkan algoritme *Machine Learning* untuk memprediksi faktor-faktor yang berperan dalam gangguan tidur. Sumber data berasal dari hasil survei mengenai pola tidur yang dilakukan oleh *National Sleep Foundation*.

Tahap awal penelitian mencakup proses pra-pemrosesan data, antara lain menghapus atau memperbaiki nilai hilang, mengatasi keberadaan outlier, melakukan transformasi menggunakan teknik encoding, serta menormalkan data agar siap dipakai dalam model. Setelah itu, data dibagi menjadi dua subset: 80% untuk pelatihan dan 20% untuk pengujian.



Gambar 1 Flowchart Metode

Dua model klasifikasi digunakan pada analisis ini, yakni *Logistic Regression* dan *XGBoost (Extreme Gradient Boosting)*. Kinerja keduanya dibandingkan dengan indikator evaluasi berupa *Accuracy*, *Precision*, *Recall*, dan *F1-score*. Selain itu, untuk menilai peran relatif setiap variabel prediktor, penelitian ini memanfaatkan metode interpretabilitas *SHAP (SHAPley Additive Explanations)*.

1. Sumber dan Pengumpulan Data

Dalam penelitian ini, data diambil dari *National Sleep Foundation (NSF)*, yaitu organisasi internasional yang bergerak di bidang penelitian, edukasi, serta advokasi terkait kesehatan tidur. Dataset yang digunakan mencakup informasi mengenai berbagai faktor yang berhubungan dengan kualitas tidur, antara lain:

TABLE 1
DATA NATIONAL SLEEP FOUNDATION

Kondisi	Tidur	Kafein	Aktivitas	Kesehatan	Usia	Kelamin
Stres Rendah	Tidak Teratur	Sering	Aktif	Sehat	Dewasa	Perempuan
Stres Tinggi	Tidak Teratur	Jarang	Kurang Aktif	Terganggu	Remaja	Perempuan
Stres Tinggi	Tidak Teratur	Jarang	Kurang Aktif	Terganggu	Dewasa	Laki
Stres Rendah	Teratur	Jarang	Kurang Aktif	Sehat	Dewasa	Perempuan
Stres Tinggi	Tidak Teratur	Sering	Aktif	Terganggu	Dewasa	Laki
Stres Rendah	Tidak Teratur	Jarang	Kurang Aktif	Sehat	Remaja	Perempuan
Stres Tinggi	Tidak Teratur	Sering	Aktif	Terganggu	Remaja	Perempuan
Stres Rendah	Teratur	Sering	Kurang Aktif	Sehat	Dewasa	Laki
Stres Rendah	Tidak Teratur	Jarang	Kurang Aktif	Sehat	Remaja	Laki
Stres Rendah	Teratur	Sering	Aktif	Sehat	Remaja	Laki
Stres Rendah	Teratur	Jarang	Kurang Aktif	Sehat	Dewasa	Laki
Stres Rendah	Tidak Teratur	Jarang	Kurang Aktif	Sehat	Remaja	Laki
Stres Rendah	Tidak Teratur	Sering	Kurang Aktif	Sehat	Remaja	Perempuan
Stres Rendah	Teratur	Sering	Aktif	Sehat	Dewasa	Laki
Stres Rendah	Teratur	Sering	Kurang Aktif	Sehat	Remaja	Laki
Stres Rendah	Tidak Teratur	Sering	Aktif	Sehat	Dewasa	Laki
Stres Rendah	Tidak Teratur	Sering	Aktif	Sehat	Remaja	Perempuan
Stres Rendah	Tidak Teratur	Sering	Aktif	Sehat	Remaja	Perempuan
Stres Tinggi	Tidak Teratur	Jarang	Aktif	Terganggu	Dewasa	Laki
Stres Rendah	Teratur	Sering	Kurang Aktif	Sehat	Dewasa	Laki
Stres Rendah	Tidak Teratur	Jarang	Kurang Aktif	Sehat	Dewasa	Perempuan
Stres Tinggi	Tidak Teratur	Jarang	Aktif	Terganggu	Dewasa	Laki
Stres Rendah	Teratur	Sering	Aktif	Sehat	Remaja	Laki
Stres Tinggi	Tidak Teratur	Sering	Aktif	Terganggu	Remaja	Laki
Stres Rendah	Teratur	Sering	Aktif	Sehat	Dewasa	Perempuan
Stres Tinggi	Teratur	Jarang	Kurang Aktif	Terganggu	Remaja	Perempuan
Stres Rendah	Teratur	Sering	Aktif	Sehat	Remaja	Laki

Dataset yang digunakan dalam penelitian ini terdiri dari 500 responden, yang bersumber dari survei berskala besar yang diselenggarakan oleh National Sleep Foundation (NSF), yang secara rutin memantau pola dan kebiasaan tidur masyarakat. Sebelum dianalisis, data tersebut telah melalui proses pembersihan (data cleansing), meliputi penanganan data yang hilang, normalisasi pada variabel numerik, serta transformasi variabel kategorikal. Tahapan ini dilakukan untuk memastikan bahwa dataset memiliki kualitas yang layak

sehingga dapat diolah dengan metode Machine Learning dalam mengungkap faktor-faktor paling berpengaruh terhadap munculnya gangguan tidur.

2. Pra-pemrosesan Data

Sebelum model dilatih, data harus dipersiapkan terlebih dahulu agar memiliki kualitas yang optimal. Tahapan pra-pemrosesan mencakup:

Pembersihan Data Menghapus entri yang kosong serta mengidentifikasi dan mengatasi nilai ekstrem (outlier) agar tidak menimbulkan bias pada hasil analisis.

Transformasi Data Mengonversi variabel kategorikal menjadi bentuk numerik menggunakan teknik label encoding sehingga dapat diproses oleh algoritma *Machine Learning*.

Normalisasi/Standarisasi Menyeragamkan skala pada variabel numerik untuk menjaga kestabilan kinerja model selama proses pelatihan, jika diperlukan.

3. Pembagian Data

Dataset dipisahkan ke dalam dua bagian utama:

- Data Training (80%): Berfungsi untuk melatih model dalam mengenali serta mempelajari pola yang terdapat pada data.
- Data Tesing (20%): Dimanfaatkan untuk mengevaluasi kinerja model setelah proses pelatihan selesai dilakukan.

4. Penerapan Algoritma

Penelitian ini menggunakan dua algoritma utama, yaitu *Logistic Regression* dan *Extreme Gradient Boosting (XGBoost)*. *Logistic Regression* digunakan untuk memodelkan hubungan linier antara variabel independen (misalnya tingkat stres, pola tidur, dan konsumsi kafein) dengan probabilitas terjadinya gangguan tidur, sehingga sesuai untuk klasifikasi biner.

Sementara itu, *XGBoost* diterapkan sebagai metode *ensemble learning* berbasis boosting yang mengombinasikan pohon keputusan secara bertahap guna memperbaiki kesalahan prediksi sebelumnya. Algoritma ini dipilih karena memiliki performa tinggi, efisiensi komputasi, serta mekanisme regularisasi yang mampu mengurangi risiko overfitting.

Kinerja model dievaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-score*. Selain itu, analisis *SHAP* digunakan untuk meningkatkan interpretabilitas model dengan menilai kontribusi setiap variabel terhadap prediksi, baik pada tingkat global maupun individual.

a. Logistic Regression

Model *Logistic Regression* memprediksi probabilitas seseorang mengalami gangguan tidur ($y=1$) berdasarkan kombinasi linier dari variabel-variabel input (X_1, X_2, \dots, X_n) [11].

$$P(y=1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}} \quad (3.1)$$

Keterangan

$P(y=1|X)$ = Probabilitas terjadinya gangguan tidur

X_1, X_2, \dots, X_n = Variabel input (misalnya stres, pola tidur, konsumsi kafein, dll.)

β_0 = Intersep, yaitu nilai dasar model saat tidak ada input

$\beta_1, \beta_2, \dots, \beta_n$ = Koefisien yang merepresentasikan besarnya pengaruh setiap variabel terhadap probabilitas gangguan tidur
e = Bilangan eksponensial (sekitar 2,718)

b. Extreme Gradient Boosting (XGBoost)

algoritma *ensemble learning* berbasis gradient boosting yang menggabungkan pohon keputusan secara bertahap. Setiap iterasi difokuskan untuk memperbaiki kesalahan prediksi sebelumnya, dengan tambahan mekanisme regularisasi agar model lebih akurat dan terhindar dari overfitting [12].

$$Obj(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^t \Omega(f_k) \quad (3.2)$$

Keterangan :

$\sum_{i=1}^n l(y_i, \hat{y}_i)$: total *loss* (kesalahan prediksi) dari seluruh data:

y_i : label sebenarnya dari data ke-iii.

$l(y_i, \hat{y}_i)$: fungsi kerugian (misalnya *logistic loss* untuk klasifikasi).

$\sum_{k=1}^t \Omega(f_k)$: penalti untuk mengontrol kompleksitas model:

f_k : pohon keputusan (*weak learner*) pada iterasi ke-k.

$\Omega(f_k)$: fungsi regularisasi (misalnya kombinasi jumlah daun dan bobot pada daun).

c. Accuracy

ukuran seberapa sering model memprediksi dengan benar dibandingkan dengan keseluruhan data yang diuji [13].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.3)$$

d. Precision

mengukur ketepatan model saat memprediksi data positif [14].

$$Precision = \frac{TP}{TP + FP} \quad (3.4)$$

Keterangan:

- Menghitung berapa banyak dari prediksi positif yang benar-benar positif.
- Semakin tinggi *Precision*, semakin sedikit kesalahan memprediksi negatif sebagai positif.

e. Recall

Mengukur kemampuan model menemukan semua data positif [15].

$$Recall = \frac{TP}{TP + FN} \quad (3.5)$$

Keterangan:

- Mengukur seberapa banyak data positif yang sebenarnya berhasil diprediksi benar.
- Semakin tinggi *Recall*, semakin sedikit kasus positif yang terlewat.

f. F1-score

rata-rata harmonis antara *Precision* dan *Recall*.

$$Recall = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.6)$$

Keterangan:

- Merupakan rata-rata harmonis antara *Precision* dan *Recall*.

- Nilainya tinggi jika *Precision* dan *Recall* keduanya tinggi.

g. Analisis SHAP

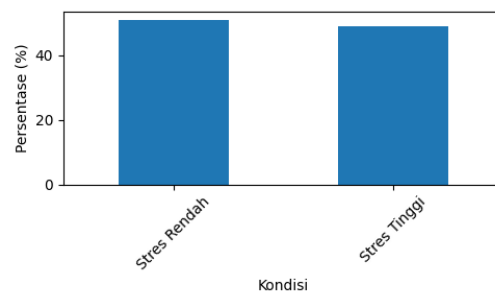
Interpretabilitas model dilakukan menggunakan SHAP (*SHAPley Additive Explanations*), yaitu metode berbasis teori game theory yang mengukur kontribusi setiap variabel terhadap hasil prediksi. SHAP digunakan untuk menilai pengaruh global melalui feature importance plot, serta menjelaskan arah kontribusi variabel pada level individu dengan decision plot [16]. Dengan demikian, hasil model tidak hanya menunjukkan tingkat akurasi, tetapi juga dapat dipahami faktor-faktor mana yang paling menentukan risiko gangguan tidur [17].

IV. IMPLEMENTASI DAN PEMBAHASAN

Bagian ini menyajikan hasil analisis data dari penelitian yang dilakukan terhadap berbagai faktor yang diduga berpengaruh terhadap kualitas tidur responden. Faktor-faktor yang dianalisis meliputi kondisi (stres), pola tidur, konsumsi kafein, aktivitas fisik, kondisi kesehatan, usia, jenis kelamin, dan gangguan tidur. Mengacu pada rekomendasi dari Sleep Foundation (*NSF*), faktor-faktor tersebut merupakan determinan penting yang dapat memengaruhi durasi, efisiensi, dan kualitas tidur seseorang. Hasil yang diperoleh kemudian dibahas untuk memahami hubungan antara masing-masing faktor dengan kualitas tidur responden secara komprehensif.

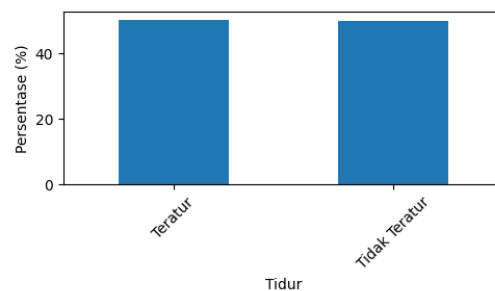
1) Kondisi (Stres)

Gambar 2 Menunjukkan Sebanyak 51,0% responden mengalami tingkat stres rendah, sementara 49,0% mengalami tingkat stres tinggi. *Persentase* ini mengindikasikan bahwa hampir setengah dari responden menghadapi tekanan psikologis yang cukup signifikan. Stres dapat mengganggu ritme sirkadian melalui peningkatan hormon kortisol yang memengaruhi regulasi tidur. Temuan ini konsisten dengan penelitian sebelumnya yang menyatakan bahwa stres kronis merupakan salah satu penyebab utama insomnia dan gangguan tidur lainnya.



Gambar 2 Kondisi (Stres)

2) Pola Tidur

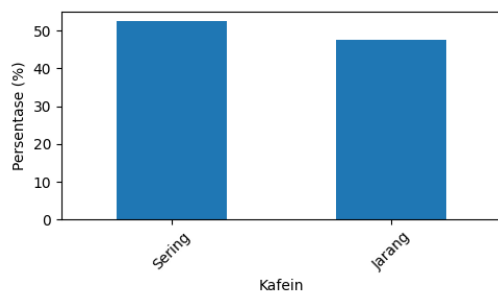


Gambar 3 Pola Tidur

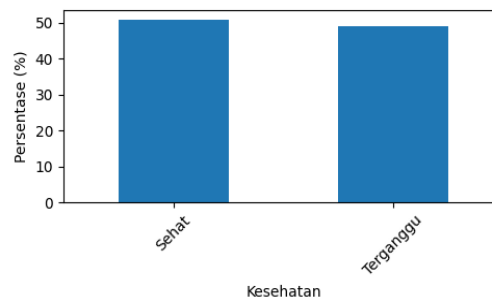
Hasil analisis Gambar 3 menunjukkan bahwa 50,5% responden memiliki pola tidur teratur, sedangkan 49,5% memiliki pola tidur tidak teratur. Pola tidur yang tidak konsisten dapat mengganggu ritme biologis tubuh, sehingga memengaruhi kualitas tidur jangka panjang. Sejumlah penelitian menyebutkan bahwa keteraturan waktu tidur dan bangun sangat penting dalam menjaga kesehatan tidur, termasuk efisiensi dan durasi tidur yang optimal.

3) Konsumsi Kafein

Gambar 4 menunjukkan sebagian besar responden, yaitu 52,2%, dilaporkan sering mengonsumsi kafein, sedangkan 47,8% jarang mengonsumsi kafein. Konsumsi kafein diketahui memiliki efek stimulasi pada sistem saraf pusat yang dapat menunda timbulnya rasa kantuk. Dengan demikian, responden yang sering mengonsumsi kafein berpotensi lebih tinggi mengalami gangguan tidur. Penelitian terdahulu juga menyebutkan bahwa konsumsi kafein berlebihan terutama menjelang waktu tidur dapat mengganggu ritme sirkadian dan mengurangi durasi tidur efektif.

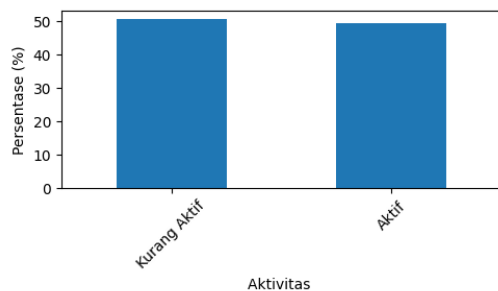


Gambar 4 Konsumsi Kafein



Gambar 6 Kesehatan

4) Aktivitas



Gambar 5 Aktivitas

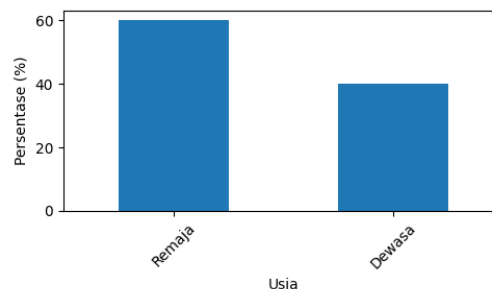
Hasil penelitian Gambar 5 menunjukkan bahwa 50,4% responden tergolong dalam kategori kurang aktif, sedangkan 49,6% tergolong aktif. Persentase ini memperlihatkan bahwa mayoritas responden cenderung memiliki tingkat aktivitas fisik yang rendah. Aktivitas fisik yang tidak optimal berpotensi mengganggu kualitas tidur karena tubuh tidak mengalami kelelahan fisik yang cukup untuk merangsang mekanisme tidur alami. Sejalan dengan penelitian sebelumnya, aktivitas fisik yang teratur terbukti dapat meningkatkan kualitas tidur melalui perbaikan sirkulasi darah dan penurunan hormon stres yang berlebihan.

5) Kesehatan

Gambar 6 Menunjukkan Responden dengan kondisi kesehatan baik tercatat sebesar 51,1%, sedangkan yang mengalami gangguan kesehatan sebesar 48,9%. Perbedaan ini meskipun relatif kecil, tetap menunjukkan kecenderungan bahwa kesehatan fisik berkorelasi dengan kualitas tidur. Individu dengan masalah kesehatan cenderung memiliki kualitas tidur yang lebih buruk karena gangguan fisik dapat menimbulkan ketidaknyamanan, nyeri, atau gangguan metabolisme yang memengaruhi siklus tidur.

6) Usia

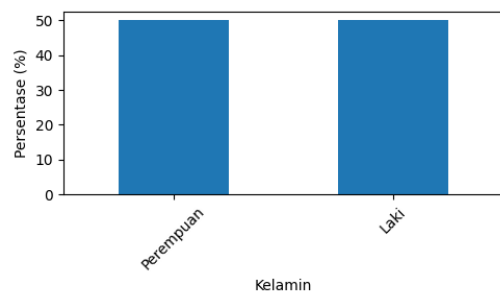
Gambar 7 Menunjukkan responden berdasarkan usia menunjukkan bahwa mayoritas berasal dari kelompok remaja, yaitu 60,2%, sedangkan dewasa sebesar 39,8%. Usia remaja sering dikaitkan dengan kebiasaan tidur yang kurang teratur, penggunaan gawai hingga larut malam, serta gaya hidup yang memicu gangguan tidur. Kondisi ini dapat menjelaskan mengapa gangguan tidur lebih banyak ditemukan pada kelompok usia muda dibandingkan kelompok dewasa.



Gambar 7 Usia

7) Jenis Kelamin

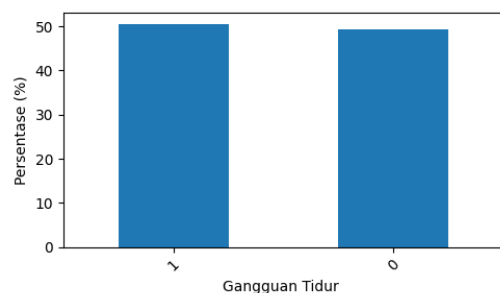
Gambar 8 Menunjukkan Data menunjukkan jenis kelamin responden hampir seimbang, dengan perempuan sebesar 50,2% dan laki-laki sebesar 49,8%. Meskipun demikian, studi-studi sebelumnya menyebutkan bahwa perempuan memiliki kerentanan yang lebih tinggi terhadap gangguan tidur, termasuk insomnia, dibandingkan laki-laki. Faktor hormonal dan psikologis diduga turut berperan dalam perbedaan ini. Analisis lebih lanjut diperlukan untuk melihat apakah perbedaan gender pada sampel penelitian ini berpengaruh signifikan terhadap kualitas tidur.



Gambar 8 Jenis Kelamin

8) Gangguan Tidur

Gambar 9 Menunjukkan Sebanyak 50,6% responden dilaporkan mengalami gangguan tidur, sementara 49,4% tidak mengalami gangguan tidur. Proporsi yang hampir seimbang ini menegaskan bahwa gangguan tidur merupakan fenomena yang cukup sering terjadi di masyarakat. Berbagai faktor, termasuk stres, pola hidup, dan kondisi kesehatan, diketahui dapat memengaruhi kualitas tidur. Temuan ini mendukung literatur sebelumnya yang menyebutkan bahwa gangguan tidur memiliki prevalensi tinggi di berbagai kelompok usia dan dapat memengaruhi produktivitas serta kesehatan mental individu.



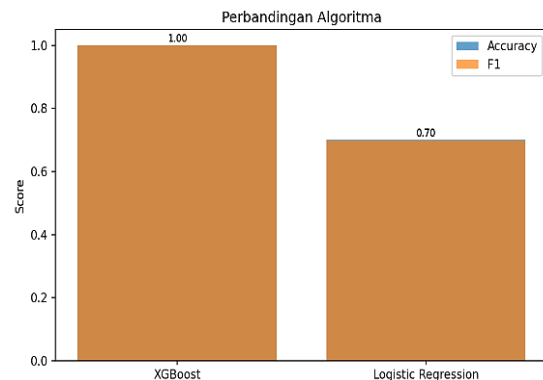
Gambar 9 Gangguan Tidur

V. PENGUJIAN

Distribusi prediksi yang relatif seimbang antara kedua kategori menunjukkan bahwa model tidak mengalami bias signifikan terhadap salah satu kelas. Dalam konteks klasifikasi biner, keseimbangan ini menjadi indikasi awal bahwa model mampu mengenali pola dari kedua kelas dengan proporsi yang hampir setara, sehingga berpotensi menghasilkan performa yang baik pada metrik evaluasi seperti *Precision* dan *Recall*, terutama jika distribusi data asli juga bersifat seimbang.

Untuk memahami kontribusi masing-masing variabel dalam pembentukan prediksi, analisis kemudian dilengkapi dengan pendekatan *SHAP* (*SHAP*ley Additive exPlanations) melalui visualisasi *SHAP*_decision, *SHAP*_importance, dan *SHAP*_summary.

1. Perbandingan Algoritma



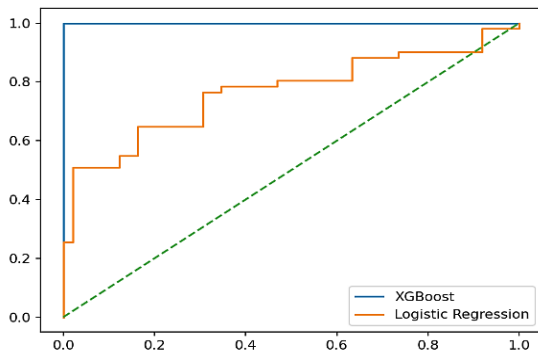
Gambar 10 Perbandingan Algoritma

Berdasarkan hasil perbandingan algoritma pada Gambar 10 terlihat bahwa algoritma *XGBoost* memperoleh nilai *Accuracy* dan *F1-score* sebesar 1.00 atau 100%, sedangkan *Logistic Regression* hanya mencapai nilai 0.70 atau 70% pada kedua metrik tersebut. Nilai *F1-score* yang setara dengan *Accuracy* menunjukkan bahwa data uji memiliki distribusi kelas yang relatif seimbang, sehingga kedua metrik memberikan hasil yang konsisten. Perbedaan skor yang signifikan ini mengindikasikan bahwa

XGBoost mampu mempelajari pola data dengan lebih baik, sehingga menghasilkan prediksi yang lebih akurat dan seimbang antara presisi dan *Recall*. Sebaliknya, kinerja

Logistic Regression yang lebih rendah mencerminkan keterbatasannya dalam menangkap hubungan non-linear atau kompleksitas data yang lebih tinggi. Hasil ini menegaskan bahwa *XGBoost* lebih unggul untuk digunakan dalam pemodelan dengan kompleksitas data yang serupa.

2. ROC Curve

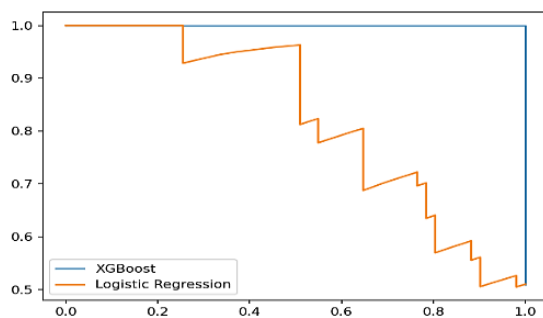


Gambar 11 ROC Curve

Gambar 11 menunjukkan kurva ROC untuk algoritma *XGBoost* dan *Logistic Regression*. Terlihat bahwa *XGBoost* memiliki kurva yang menempel di sisi kiri atas grafik dengan $AUC = 1.00$, menandakan kinerja klasifikasi yang hampir sempurna. Sementara itu, *Logistic*

Regression memiliki $AUC = 0.85$, yang menunjukkan kinerja cukup baik tetapi masih di bawah *XGBoost*. Semakin luas area di bawah kurva, semakin baik kemampuan model dalam membedakan kelas positif dan negatif. Hasil ini mengonfirmasi bahwa *XGBoost* lebih unggul dalam memprediksi dibandingkan *Logistic Regression*.

3. Precision-Recall (PR) Curve



Gambar 12 Precision-Recall Curve

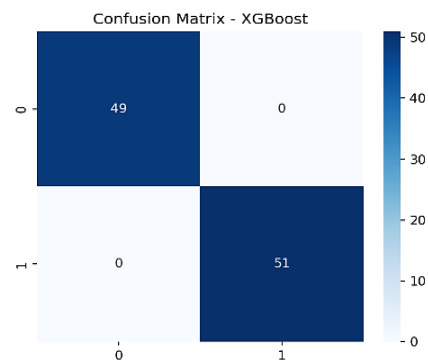
Gambar 12 menunjukkan kurva *Precision-Recall* untuk algoritma *XGBoost* dan *Logistic Regression*. Terlihat bahwa *XGBoost* memiliki kurva yang mendekati area atas grafik dengan nilai $AUC-PR = 1.00$, menunjukkan presisi dan *Recall* yang tinggi sekaligus seimbang. Sementara itu, *Logistic*

Regression memiliki $AUC-PR = 0.75$, yang mengindikasikan kinerja lebih rendah dalam

mendeteksi kelas positif secara akurat. Secara keseluruhan, hasil ini menegaskan bahwa *XGBoost* lebih unggul dalam mempertahankan keseimbangan antara presisi dan *Recall* dibandingkan *Logistic Regression*.

4. Perbandingan Analisis Confusion Matrix

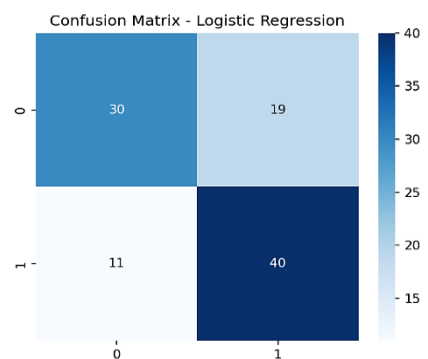
a. Confusion Matrix Gradient Boosting



Gambar 13 Confusion Matrix XGBoost

Gambar 13 menunjukkan *Confusion Matrix* untuk algoritma *XGBoost*, di mana dari total 100 data uji, model berhasil mengklasifikasikan 49 sampel kelas negatif dan 51 sampel kelas positif dengan benar tanpa kesalahan klasifikasi ($True\ Positive = 51$, $True\ Negative = 49$, $False\ Positive = 0$, $False\ Negative = 0$). Hasil ini menghasilkan akurasi, presisi, *Recall*, dan *F1-score* sebesar 1.00, menandakan bahwa *XGBoost* memiliki kinerja sempurna pada dataset ini.

b. Confusion Matrix Logistic Regression



Gambar 14 Confusion Matrix Logistic Regression

Gambar 14 menunjukkan *Confusion Matrix* untuk algoritma *Logistic Regression*, di mana

dari total 100 data uji, model mengklasifikasikan 30 sampel kelas negatif dan 40 sampel kelas positif dengan benar.

Namun, terdapat 19 kesalahan prediksi positif palsu (*False Positive*) dan 11 kesalahan negatif palsu (*False Negative*). Berdasarkan hasil ini, diperoleh akurasi 70%, presisi 67.8%, *Recall* 78.4%, dan *F1-score* 72.7%, yang menunjukkan bahwa *Logistic Regression* memiliki kinerja moderat dengan tingkat kesalahan yang cukup tinggi dibandingkan *XGBoost*.

5. Evaluasi

Model	Accuracy	Precision	Recall	F1-Score
XGBoost	1.00	1.00	1.00	1.00
Logistic Regression	0.70	0.7043	0.70	0.6976

Gambar 15 Evaluasi

a. Accuracy

- Mengukur proporsi prediksi yang benar terhadap seluruh data uji.
- *XGBoost* = 1.0 → Semua prediksi tepat (100%).
- *Logistic Regression* = 0.7 → Hanya 70% prediksi yang benar, sedangkan 30% sisanya salah klasifikasi.

b. Precision

- Menggambarkan ketepatan klasifikasi positif: dari seluruh data yang diprediksi positif, berapa yang benar-benar positif.
- *XGBoost* = 1.0 → Tidak ada *False Positive*, semua prediksi positif benar.
- *Logistic Regression* = 0.7043 → Sekitar 30% prediksi positif keliru (*False Positive*).

c. Recall

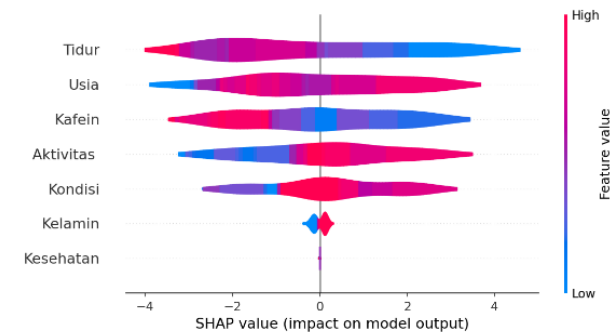
- Mengukur kemampuan model dalam mendeteksi seluruh data positif: dari semua data positif, berapa yang berhasil terdeteksi.
- *XGBoost* = 1.0 → Semua data positif berhasil terdeteksi (100%).
- *Logistic Regression* = 0.7 → Hanya 70% data positif terdeteksi, sisanya terlewat sebagai *False Negative*.

d. F1-score

- Rata-rata harmonik antara *Precision* dan *Recall*, memberikan gambaran keseimbangan antara keduanya.

- *XGBoost* = 1.0 → Keseimbangan sempurna karena *Precision* dan *Recall* sama-sama 1.0.
- *Logistic Regression* = 0.698 → Menunjukkan ketidakseimbangan karena *Precision* dan *Recall* sama-sama rendah.

6. SHAP Summary (Violin) Plot.



Gambar 16 SHAP Violin Plot

Gambar 16 tersebut merupakan *plot summary SHAP (SHAPley Additive exPlanations)* yang memperlihatkan kontribusi variabel terhadap output model. Sumbu horizontal menunjukkan nilai *SHAP* sebagai dampak fitur terhadap prediksi, sedangkan gradasi warna biru–merah merepresentasikan nilai fitur dari rendah hingga tinggi.

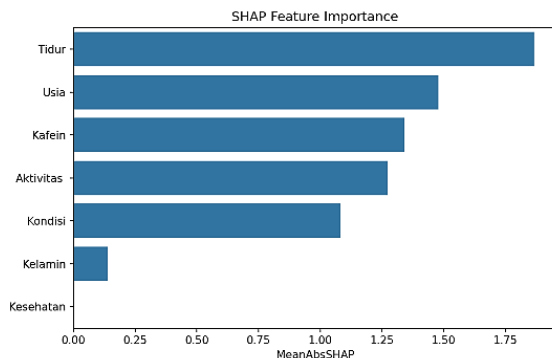
Terlihat bahwa *Tidur*, *Usia*, *Kafein*, dan *Aktivitas* memiliki pengaruh dominan dengan variasi nilai yang jelas, sementara *Kondisi* berpengaruh sedang, serta *Kelamin* dan *Kesehatan* menunjukkan kontribusi yang relatif kecil atau hampir tidak berpengaruh terhadap model.

7. SHAP Feature Importance bar plot

Gambar 17 tersebut menunjukkan hasil analisis *SHAP Feature Importance* yang merepresentasikan kontribusi rata-rata absolut nilai *SHAP* setiap variabel terhadap model. Variabel *Tidur* menempati posisi paling dominan dengan nilai sekitar 1,9, menunjukkan bahwa durasi atau kualitas tidur merupakan faktor utama yang memengaruhi prediksi gangguan tidur.

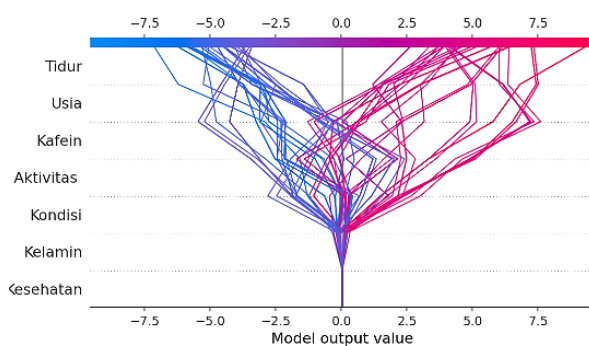
Selanjutnya, *Usia* ($\pm 1,5$), *Kafein* ($\pm 1,3$), dan *Aktivitas* ($\pm 1,25$) juga berperan penting, mengindikasikan bahwa faktor demografis dan gaya hidup turut memberikan kontribusi signifikan. Sementara itu, *Kondisi* kesehatan umum memiliki pengaruh sedang ($\pm 1,1$),

sedangkan Kelamin ($\pm 0,15$) dan Kesehatan ($\pm 0,05$) hanya memberikan kontribusi minimal, yang menunjukkan bahwa variabel tersebut relatif kurang berperan dalam model.



Gambar 17 SHAP Importance bar plot

8. SHAP decision plot



Gambar 18 SHAP decision plot

Gambar 18 tersebut menunjukkan SHAP decision plot yang menggambarkan kontribusi kumulatif setiap variabel terhadap output model. Terlihat bahwa variabel Tidur memberikan pengaruh terbesar dengan rentang nilai sekitar $-7,5$ hingga $+8,0$ terhadap prediksi, diikuti oleh Usia ($-6,5$ hingga $+6,0$), Kafein ($-5,5$ hingga $+5,0$), dan Aktivitas ($-5,0$ hingga $+4,5$). Sementara itu, Kondisi berkontribusi sedang (sekitar $-3,5$ hingga $+3,0$), sedangkan Kelamin ($-1,0$ hingga $+1,0$) dan Kesehatan (mendekati 0) menunjukkan pengaruh yang relatif kecil terhadap hasil model.

I. KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan algoritma *Machine Learning*, khususnya *XGBoost*, memiliki performa jauh lebih unggul dibandingkan *Logistic Regression* dalam memprediksi faktor-faktor yang

berkontribusi terhadap gangguan tidur. Hasil evaluasi menunjukkan bahwa *XGBoost* mencapai nilai akurasi, presisi, *Recall*, dan *F1-score* sebesar 100%, sedangkan *Logistic Regression* hanya mencapai sekitar 70%. Analisis interpretabilitas menggunakan SHAP mengungkapkan bahwa faktor utama yang memengaruhi prediksi gangguan tidur adalah pola tidur ($\pm 1,9$), usia ($\pm 1,5$), konsumsi kafein ($\pm 1,3$), dan aktivitas fisik ($\pm 1,25$). Sementara itu, faktor kondisi kesehatan umum berpengaruh sedang ($\pm 1,1$), sedangkan jenis kelamin ($\pm 0,15$) dan kesehatan subjektif ($\pm 0,05$) memiliki kontribusi minimal.

Temuan ini menegaskan bahwa gaya hidup dan kebiasaan harian merupakan determinan dominan kualitas tidur, sejalan dengan literatur terdahulu. Selain itu, penggunaan SHAP terbukti memberikan interpretasi yang transparan terhadap model sehingga memudahkan pemahaman praktis atas hasil prediksi.

Ke depan, pengembangan dapat diarahkan pada penggunaan dataset yang lebih luas dan beragam, integrasi dengan data fisiologis *real-time* (misalnya dari *wearable device*), serta penerapan model *ensemble learning* berbasis deep learning untuk meningkatkan generalisasi. Selain itu, sistem ini dapat diperluas menjadi aplikasi prediksi mandiri yang mampu memberikan rekomendasi personal terkait pola tidur sehat dan strategi pencegahan gangguan tidur.

REFERENSI

- [1] Mohammad Rasel Mahmud, Al Shahriar Uddin Khondakar Pranta, Anamul Haque Sakib, Abdullah Al Sakib, and Md Ismail Hossain Siddiqui, "Robust feature selection for improved sleep stage classification," *Int. J. Sci. Res. Arch.*, vol. 15, no. 1, pp. 1790–1797, 2025, doi: 10.30574/ijrsra.2025.15.1.1160.
- [2] J. P. Ntayagabiri, Y. Bentaleb, J. Ndikumagenge, and H. El Makhtoum, "A Comparative Analysis of Supervised Machine Learning Algorithms for IoT Attack Detection and Classification," *J. Comput. Theor. Appl.*, vol. 2, no. 3, pp. 395–409, 2025, doi: 10.62411/jcta.11901.

- [3] G. Long, W. Qingyan, T. Jian, W. Duo, W. Xiaotian, and D. Hongfei, "航空工程进展 Advances in Aeronautical Science and Engineering," *Adv. Aeronaut. Sci. Eng.*, vol. 14, no. 2, pp. 178–190, 2023.
- [4] S. Aziz *et al.*, "Wearable Artificial Intelligence for Sleep Disorders: Scoping Review," *J. Med. Internet Res.*, vol. 27, 2025, doi: 10.2196/65272.
- [5] S. Khanmohmmadi, T. Khatibi, G. Tajeddin, E. Akhondzadeh, and A. Shojaee, "Revolutionizing sleep disorder diagnosis: A Multi-Task learning approach optimized with genetic and Q-Learning techniques," *Sci. Rep.*, vol. 15, no. 1, pp. 1–16, 2025, doi: 10.1038/s41598-025-01893-4.
- [6] A. D. Septiadi, M. Awiet, and W. Prasetyo, "Clustering Sleep Patterns and Health Metrics Using K-Means Algorithm to Identify Profiles of Sleep Quality and Well-being in a Diverse Population," vol. 5, no. 2, pp. 125–138, 2025.
- [7] P. Guyot, M. Eveilleau, T. Bastogne, C. Ayav, N. Carpentier, and B. Chenuel, "Robust performances of a nocturnal long-term ECG algorithm for the evaluation of sleep apnea syndrome: A pilot study," *PLoS One*, vol. 20, no. 5 May, pp. 1–15, 2025, doi: 10.1371/journal.pone.0318622.
- [8] M. Mostafa Monowar *et al.*, "Advanced sleep disorder detection using multi-layered ensemble learning and advanced data balancing techniques," *Front. Artif. Intell.*, vol. 7, 2024, doi: 10.3389/frai.2024.1506770.
- [9] Muhammad Iqbal, Iin Aini Isnawati, and Nur Hamim, "Hubungan Kualitas Tidur Dengan Kesehatan Mental Remaja Yang Menggunakan Gadget Di Smkn 1 Banyuanyar," *J. Cent. Publ.*, vol. 3, no. 3, pp. 241–251, 2025, doi: 10.60145/jcp.v3i3.370.
- [10] N. Khasanah, D. U. Eka Saputri, F. Aziz, and T. Hidayat, "Studi Perbandingan Algoritma Random Forest dan K-Nearest Neighbors (KNN) dalam Klasifikasi Gangguan Tidur," *Comput. Sci.*, vol. 5, no. 1, pp. 17–25, 2025, doi: 10.31294/coscience.v5i1.5522.
- [11] M. Jain and A. Srihari, "Comparison of Machine Learning Algorithm in Intrusion Detection Systems: A Review Using Binary Logistic Regression," *Int. J. Comput. Sci. Mob. Comput.*, vol. 13, no. 10, pp. 45–53, 2024, doi: 10.47760/ijcsmc.2024.v13i10.005.
- [12] M. Wiens, A. Verone-Boyle, N. Henscheid, J. T. Podichetty, and J. Burton, "A Tutorial and Use Case Example of the eXtreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications," *Clin. Transl. Sci.*, vol. 18, no. 3, 2025, doi: 10.1111/cts.70172.
- [13] M. Alshamrani, Sager Alharthi, M. Helmi, and T. Alwadei, "Journal of Business and Management Studies Determinants of Employee Retention in Pharmaceutical Companies: Case of Saudi Arabia," *Determ. Empl. Retent. Pharm. Co. Case Saudi Arab.*, no. 2709–0876, pp. 8–22, 2023, doi: 10.32996/jbms.
- [14] A. R. Zarkasi and F. R. Doni, "Analisis Data Produksi Menggunakan Metode Decision Tree (C4.5) di PT.Baja Marga Kharisma Utama (BMKU)," *J. Inform. dan Multimed.*, vol. 17, no. 1, pp. 28–37, 2025, doi: 10.33795/jtim.v17i1.6429.
- [15] I. Recognition *et al.*, "SCIENTIFIC INSIGHTS AND IMAGE RECOGNITION BASED MULTI PATH RECALL AND," pp. 11–20, 2025.
- [16] S. M. K. Alam, P. Li, M. Rahman, M. Fida, and V. Elumalai, "Key factors affecting groundwater nitrate levels in the Yinchuan Region, Northwest China: Research using the eXtreme Gradient Boosting (XGBoost) model with the SHapley Additive exPlanations (SHAP) method," *Environ. Pollut.*, vol. 364, no. P1, p. 125336, 2025, doi: 10.1016/j.envpol.2024.125336.
- [17] S. S. Chandra, R. Kumar, A. Arjunasamy, S. Galagali, A. Tantri, and S. R. Naganna, "Predicting

the compressive strength of polymer-infused bricks: A machine learning approach with SHAP interpretability,”

Sci. Rep., vol. 15, no. 1, pp. 1–22, 2025, doi: 10.1038/s41598-025-89606-9.