

# Early Detection of Sleep Disorders Using Ensemble Boosting and Classical ML with Lifestyle Data

Muhammad Rayhan Ramadhani<sup>1\*</sup>, Sendi Novianto<sup>2\*</sup>

Teknik Informatika, Universitas Dian Nuswantoro  
[111202113781@mhs.dinus.ac.id](mailto:111202113781@mhs.dinus.ac.id)<sup>1</sup>, [sendi.novianto@dsn.dinus.ac.id](mailto:sendi.novianto@dsn.dinus.ac.id)<sup>2</sup>

## Article Info

### Article history:

Received 2026-01-20

Revised 2026-03-02

Accepted 2026-04-10

### Keyword:

*Sleep Disorder,  
Machine Learning,  
Ensemble Boosting,  
Classical Machine Learning,  
Predictive Modelling,  
Lifestyle Data.*

## ABSTRACT

Sleep disorders are increasingly prevalent in modern society, significantly impacting quality of life, productivity, and physical and mental health. This study distinguishes itself by evaluating six machine learning algorithms—Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbor (KNN), XGBoost, CatBoost, and LightGBM—using a comprehensive preprocessing pipeline that ensures balanced, normalized, and leakage-free training data, enabling a robust comparison of classical and ensemble boosting models for sleep disorder prediction. Data preprocessing included handling missing values, encoding categorical features, normalization, class imbalance correction via Random Oversampling, and stratified train-test splitting. Models were optimized through hyperparameter tuning with GridSearchCV and evaluated using accuracy, precision, recall, F1-score, and multiclass ROC-AUC metrics. Feature importance analysis revealed that Age, Occupation, Heart Rate, Sleep Duration, and Daily Steps were the most influential predictors, highlighting the interpretability of the models. LightGBM achieved the highest predictive performance with 70.0% accuracy and a ROC-AUC of 0.736, followed by Naïve Bayes (67.5% accuracy, ROC-AUC 0.784) and XGBoost (65.0% accuracy, ROC-AUC 0.728). McNemar's test indicated no statistically significant difference in performance among models, confirming prediction stability. Models also differed in computational efficiency, with Naïve Bayes being the fastest and CatBoost the slowest, reflecting algorithmic complexity. These findings suggest that ensemble boosting algorithms, particularly LightGBM and XGBoost, alongside classical models like Naïve Bayes, provide effective, interpretable, and reliable tools for early detection of sleep disorder risk, with potential applications in wearable devices, digital health platforms, and clinical monitoring systems.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

## I. PENDAHULUAN

Gangguan tidur merupakan salah satu permasalahan kesehatan yang semakin banyak terjadi di masyarakat modern dan memiliki dampak besar terhadap kualitas hidup, produktivitas, serta kesehatan fisik dan mental seseorang [1]. Laporan terbaru dari World Health Organization juga menegaskan bahwa gangguan tidur berkontribusi terhadap peningkatan risiko penyakit tidak menular dan gangguan kesehatan mental secara global. Pola tidur yang buruk dapat menyebabkan penurunan konsentrasi, gangguan emosi, serta meningkatkan risiko

penyakit kronis seperti hipertensi dan gangguan kardiovaskular [2], [3]. Gangguan tidur seperti Obstructive Sleep Apnea (OSA) diketahui sebagai salah satu faktor risiko utama terjadinya hipertensi dan penyakit jantung [4], dan temuan ini masih diperkuat oleh studi epidemiologi terbaru dalam lima tahun terakhir yang menunjukkan hubungan signifikan antara sleep apnea dan risiko kardiometabolik [3], [5]

Selain itu, faktor gaya hidup seperti penggunaan gadget dan beban aktivitas harian juga terbukti berpengaruh terhadap kualitas tidur dan kesehatan mental, khususnya pada usia produktif dan remaja [6], [7]. Kualitas tidur yang

buruk juga memiliki hubungan erat dengan performa akademik serta fungsi kognitif individu [8]. Studi terbaru menunjukkan bahwa pola aktivitas harian dan tekanan kerja memiliki korelasi signifikan terhadap peningkatan risiko insomnia pada populasi usia produktif [9], [10].

Perkembangan teknologi informasi membuka peluang besar dalam pemanfaatan machine learning untuk membantu menganalisis dan memprediksi kondisi kesehatan, termasuk gangguan tidur [11], [12]. Dalam lima tahun terakhir, pendekatan machine learning berbasis data gaya hidup dan parameter kesehatan menunjukkan peningkatan akurasi dalam klasifikasi gangguan tidur [13], [14]. Integrasi faktor kesehatan dan lingkungan dengan algoritma machine learning dapat digunakan untuk memprediksi kualitas tidur secara efektif [9]. Namun, sebagian besar penelitian sebelumnya masih menggunakan satu atau dua algoritma saja, sehingga belum memberikan gambaran menyeluruh mengenai performa berbagai algoritma populer dalam kasus yang sama.

Di sisi lain, berbagai penelitian di bidang machine learning menunjukkan bahwa algoritma ensemble boosting seperti XGBoost [15], LightGBM [16], dan CatBoost [17] memiliki kemampuan yang sangat baik dalam menangani data kompleks dan tidak seimbang. Berbagai studi terbaru (2021–2024) masih menunjukkan dominasi algoritma boosting dalam klasifikasi data kesehatan berbasis tabular [18], [19], [20].

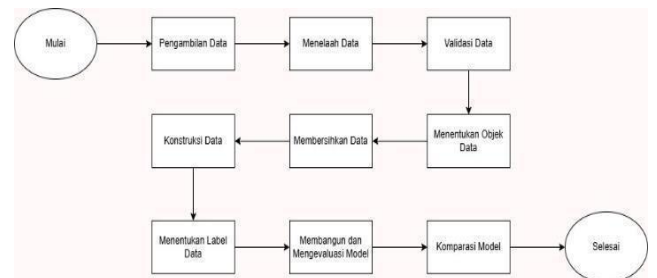
Selain itu, algoritma lain seperti Support Vector Machine (SVM), Naïve Bayes, dan K-Nearest Neighbor (KNN) juga banyak digunakan dalam klasifikasi data kesehatan karena kesederhanaan dan interpretabilitasnya [21], [22], [23]. Meskipun merupakan metode klasik, penelitian terkini masih memanfaatkan algoritma tersebut sebagai pembandingan dalam studi komparatif machine learning, namun perbandingan kinerjanya dalam satu kerangka penelitian yang sama pada kasus gangguan tidur masih jarang dilakukan.

Penelitian ini membandingkan efektivitas beberapa algoritma dalam memprediksi kemungkinan masalah tidur menggunakan pola aktivitas dan kerja sehari-hari, antara lain SVM, Naïve Bayes, KNN, XGBoost, CatBoost, dan LightGBM. Tujuan dari penelitian ini adalah untuk membangun model prediksi yang andal dan mengidentifikasi metode optimal dengan menggunakan ukuran evaluasi seperti akurasi, precision, recall, dan F1-score, serta melakukan pengujian signifikansi statistik antar model untuk memastikan validitas perbandingan performa. Diharapkan penelitian ini dapat berkontribusi dalam pengembangan sistem pendukung keputusan berbasis machine learning untuk deteksi dini risiko gangguan tidur [14].

**II. METODE PENELITIAN**

Penelitian ini menggunakan pendekatan kuantitatif eksperimental untuk mengidentifikasi potensi gangguan tidur berdasarkan pekerjaan dan pola aktivitas sehari-hari

melalui penerapan algoritma machine learning. Pendekatan kuantitatif dipilih karena memungkinkan analisis hubungan antar variabel secara objektif dan terukur. Seluruh proses pengolahan data, pemodelan, dan evaluasi dilakukan menggunakan Google Colaboratory (Google Colab), sebuah lingkungan komputasi berbasis cloud yang banyak digunakan dalam penelitian data science karena fleksibilitas, kemampuan komputasi tinggi, serta kemudahan replikasi eksperimen [24], [25]. Gambar 1 menunjukkan langkah-langkah penelitian.



Gambar 1. Tahapan penelitian

Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle, yaitu Sleep Health and Lifestyle Dataset. Dataset terbuka ini mendukung transparansi dan reproducibility penelitian machine learning [13]. Dataset terdiri dari 374 baris data dengan 13 atribut yang merepresentasikan faktor kesehatan tidur dan gaya hidup individu, termasuk usia, durasi tidur, tingkat aktivitas fisik, tingkat stres, tekanan darah, kategori BMI, dan jenis pekerjaan. Variabel target adalah Sleep Disorder, yang digunakan untuk mengklasifikasikan kondisi gangguan tidur individu. Fitur-fitur dari dataset yang digunakan beserta keterangannya tercantum pada Tabel I.

TABEL I  
FITUR DATASE

Fitur	Tipe Data	Nilai	Keterangan
Person ID	Int64	1 – 374	Identitas Unik untuk Setiap Individu Dalam Dataset
Gender	Object	Male and Female	Jenis Kelamin Biologis Individu
Age	Int64	27 – 59	Usia Individu
Occupation	Object	11 Jenis Pekerjaan	Jenis Pekerjaan Individu
Sleep Duration	Float64	5,8 – 8,5 Jam	Rata-Rata Jam Tidur Individu
Quality of Sleep	Int64	4 – 9 (skala 1 – 10)	Qualitas Tidur Pada Individu
Pyshical Activity Level	Int64	30 – 90 Menit	Rata-Rata Aktivitas Fisik Berat Harian yang Diukur Dalam Menit
Stress Level	Int64	3 – 8 (Skala 1 – 10)	Tingkat stres pada individu

BMI Category	Object	Normal, Overweight, Obese	Kategori BMI Individu
Blood Pressure	Object	115 – 142/ 75 – 95	Tekanan Darah Dalam mmhg (sistolik/diastolik)
Heart Rate	Int64	65 – 86	Detak Jantung (bpm)
Daily Steps	Int64	3000 – 10000	Jumlah Langkah Pada Individu
Sleep Disorder	Object	None, Sleep Apnea, Insomnia	Kondisi Gangguan Tidur Pada Individu

Tahap awal pengolahan data dilakukan melalui eksplorasi dataset untuk memahami struktur data, tipe atribut, jumlah data, serta mendeteksi nilai yang hilang, inkonsistensi, dan duplikasi yang dapat memengaruhi kualitas prediksi [26], [27]. Nilai NaN pada atribut Sleep Disorder dikategorikan ulang sebagai “Normal” agar variabel target dapat digunakan secara optimal dalam klasifikasi multikelas. Seluruh data juga diperiksa untuk mendeteksi duplikasi dan inkonsistensi, termasuk pengecekan outlier pada fitur numerik. Selanjutnya, atribut dipisahkan menjadi numerik dan kategorikal. Fitur kategorikal seperti Gender, Occupation, Physical Activity Level, BMI Category, dan Sleep Disorder dikonversi ke bentuk numerik menggunakan Label Encoding [22], [28]. Fitur numerik dinormalisasi menggunakan StandardScaler agar skala antar fitur seragam, penting karena beberapa algoritma, seperti Support Vector Machine dan K-Nearest Neighbor, sensitif terhadap perbedaan skala data [27], [29]. Kolom Blood Pressure yang awalnya berupa string dipecah menjadi dua kolom numerik, yaitu Systolic dan Diastolic.

Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 70:30 menggunakan stratified sampling untuk menjaga proporsi kelas target tetap seimbang. Selama pelatihan, Stratified K-Fold Cross-Validation diterapkan untuk memastikan distribusi kelas seimbang di setiap fold dan meningkatkan validitas evaluasi model [18], [30]. Jika terdapat ketidakseimbangan kelas pada variabel target, dilakukan Random Oversampling hanya pada data latih melalui pipeline, sehingga kelas minoritas diperkuat tanpa memengaruhi data uji dan mencegah data leakage [31], [32]. Distribusi kelas sebelum dan sesudah oversampling divisualisasikan untuk memastikan proses balancing berjalan efektif.

Enam algoritma klasifikasi diterapkan, yaitu K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes, XGBoost, LightGBM, dan CatBoost. Setiap algoritma dilatih melalui pipeline yang menggabungkan normalisasi, oversampling, dan model, serta dilakukan hyperparameter tuning menggunakan GridSearchCV agar performa model optimal. Hyperparameter tuning meliputi eksplorasi nilai-nilai kandidat untuk parameter penting masing-masing model, seperti jumlah tetangga (KNN), kernel dan parameter regularisasi (SVM), smoothing (Naïve Bayes), dan jumlah estimator, learning rate, serta

kedalaman pohon (XGBoost, LightGBM, CatBoost[19], [20]).

Evaluasi performa model dilakukan menggunakan metrik accuracy, precision, recall, F1-score, serta ROC-AUC multiclass untuk menilai kemampuan model dalam membedakan setiap kelas secara simultan [19], [33]. Seluruh proses evaluasi divisualisasikan melalui confusion matrix, kurva ROC-AUC, dan grafik perbandingan akurasi antar model, sehingga interpretasi performa masing-masing algoritma lebih jelas.

Waktu eksekusi setiap model dicatat untuk menilai efisiensi komputasi, yang menjadi pertimbangan penting dalam implementasi nyata pada sistem pendukung keputusan dan aplikasi berbasis wearable devices [23]. Selain itu, interpretabilitas model dianalisis menggunakan feature importance pada algoritma boosting seperti XGBoost, LightGBM, dan CatBoost untuk mengidentifikasi atribut paling berpengaruh terhadap prediksi gangguan tidur. Analisis ini penting untuk memahami kontribusi masing-masing fitur dan mendukung rekomendasi berbasis data [18], [34].

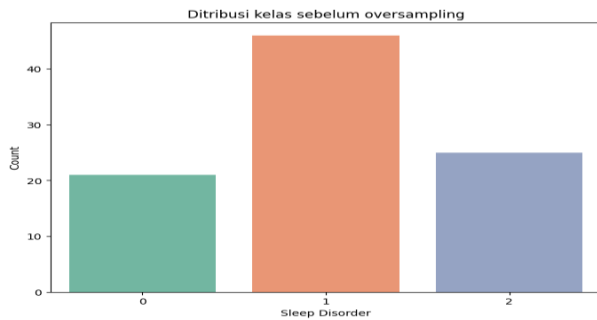
Kombinasi langkah-langkah preprocessing, hyperparameter tuning, cross-validation, evaluasi metrik, visualisasi, analisis waktu training, dan interpretabilitas model menjadi dasar metodologi penelitian ini, sehingga dapat memilih algoritma yang paling efektif dan efisien dalam memprediksi risiko gangguan tidur.

### III. HASIL DAN PEMBAHASAN

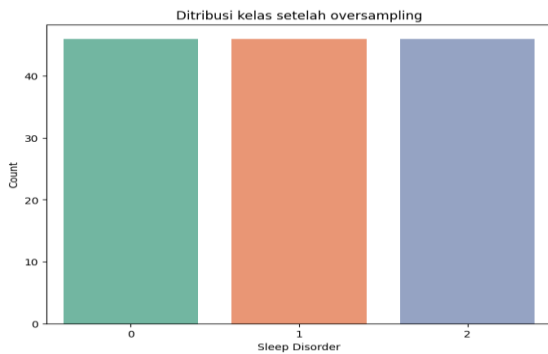
Tahap awal analisis dilakukan dengan menelaah dataset Sleep Health and Lifestyle Dataset untuk memahami karakteristik data yang digunakan. Dataset terdiri dari 374 baris data dengan 13 atribut yang mencerminkan kondisi kesehatan, aktivitas harian, dan gaya hidup individu. Penelaahan ini bertujuan untuk memastikan kesesuaian data dengan tujuan penelitian, yaitu memprediksi risiko gangguan tidur berdasarkan pola aktivitas harian dan pekerjaan. Dari hasil penelaahan, diketahui bahwa dataset mengandung atribut numerik dan kategorikal, sehingga diperlukan perlakuan pra-pemrosesan yang berbeda pada masing-masing jenis atribut. Selanjutnya dilakukan proses validasi data untuk mengidentifikasi nilai kosong, data duplikat, dan inkonsistensi data [26]. Hasil validasi menunjukkan adanya beberapa nilai kosong pada kolom Sleep Disorder, yang merepresentasikan kondisi individu normal. Nilai tersebut kemudian dikategorikan ulang sebagai “Normal” agar variabel target dapat digunakan secara optimal dalam proses klasifikasi multikelas. Selain itu, ditemukan duplikasi dan inkonsistensi pada kolom BMI Category yang kemudian ditangani agar data konsisten dan tidak memengaruhi pemodelan machine learning.

Atribut dalam dataset dipisahkan menjadi atribut numerik dan kategorikal untuk memudahkan penerapan teknik pra-pemrosesan yang sesuai [27]. Atribut kategorikal seperti Gender, Occupation, Blood Pressure,

dan BMI Category dikonversi ke bentuk numerik menggunakan Label Encoding, sementara atribut numerik dinormalisasi menggunakan StandardScaler agar algoritma seperti SVM dan KNN yang sensitif terhadap skala dapat bekerja optimal [27]. Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 70:30 menggunakan Stratified K-Fold untuk memastikan distribusi kelas tetap seimbang pada kedua subset data [32]. Penanganan ketidakseimbangan kelas pada data latih dilakukan dengan metode Random Oversampling yang diterapkan dalam pipeline pelatihan untuk menghindari data leakage [31], [32].



Gambar 2. Target Sebelum Oversampling



Gambar 3. Target Setelah Oversampling

Enam algoritma machine learning diterapkan dalam penelitian ini, yaitu Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbor (KNN), XGBoost, LightGBM, dan CatBoost. Proses hyperparameter tuning dilakukan menggunakan GridSearchCV untuk setiap algoritma dengan tujuan memperoleh konfigurasi model terbaik yang menghasilkan performa optimal [17], [35]. Hasil GridSearchCV tertera pada Tabel II.

TABEL II  
HYPERPARAMETER GRIDSEARCHCV

Model	Parameter	Nilai	Nilai Terbaik
KNN	n_neighbours, weights	(3,5,7,9), (uniform, distance)	7, uniform
SVM	C, gamma, kernel	(0.1,1,10), (scale,auto), (rbf)	1, scale, rbf
Naïve Bayes	var_smoothing	(1e-9, 1e-8, 1e-7)	1e-7

XGBoost	n_estimators, max_depth, learning_rate	(100, 200), (3, 5), (0.01, 0.1)	200, 3, 0.01
LightGBM	n_estimators, max_depth, learning_rate	(100, 200), (-1, 5, 10), (0.01, 0.1)	200, -1, 0.01
Catboost	iterations, depth, learning_rate	(200, 300), (4, 6, 8), (0.01, 0.1)	300, 4, 0.01

Parameter dan nilai terbaik masing-masing model diambil berdasarkan performa cross-validation yang paling tinggi, sehingga memastikan hasil hyperparameter tidak hanya berlaku untuk satu subset data, tetapi stabil di seluruh data latih [7], [34]. Hasil CV mean dan CV STD bisa dilihat pada Tabel III.

TABEL III  
CROSS VALIDATION MEAN & STD

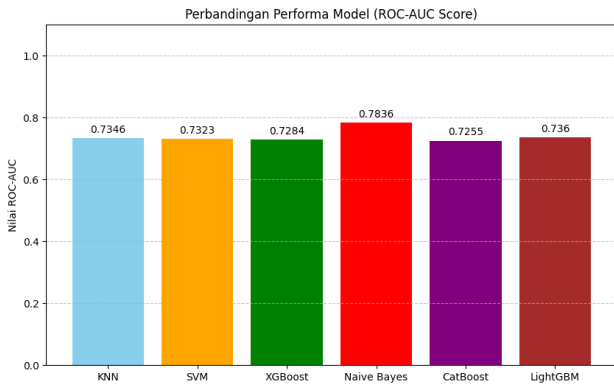
Model	CV Mean	CV STD
KNN	67%	4%
SVM	73%	9%
Naïve Bayes	74%	10%
XGBoost	74%	8%
LightGBM	74%	13%
Catboost	79%	11%

Evaluasi model dilakukan menggunakan metrik accuracy, precision, recall, F1-score, dan ROC-AUC multiclass untuk menilai kemampuan model dalam membedakan tiap kategori gangguan tidur. Selain itu, waktu pelatihan (training time) dicatat untuk membandingkan efisiensi komputasi masing-masing model [15]. Selanjutnya, feature importance dianalisis pada algoritma boosting seperti LightGBM untuk mengetahui variabel yang paling memengaruhi prediksi gangguan tidur [16]. Untuk memverifikasi perbedaan performa antar model secara statistik, dilakukan uji McNemar, yang berguna untuk menilai signifikansi perbedaan akurasi pada dataset uji yang sama [36]. Perbandingan hasil model bisa dilihat pada Tabel IV.

TABEL IV  
ACCURACY, PRECISION, RECALL, DAN F1-SCORE

Model	Precision, Recall, F1-score			Accuracy	Time
	Insomnia	Normal	Sleep Apnea		
XGBoost	50%, 56%, 53%	73%, 86%, 79%	50%, 22%, 31%	65%	5.29
CatBoost	0.5, 0.67, 0.57	0.78, 0.82, 0.8	40%, 22%, 29%	65%	21.44
Naïve Bayes	50%, 67%, 67%	78%, 82%, 80%	60%, 33%, 43%	67.5%	0.17

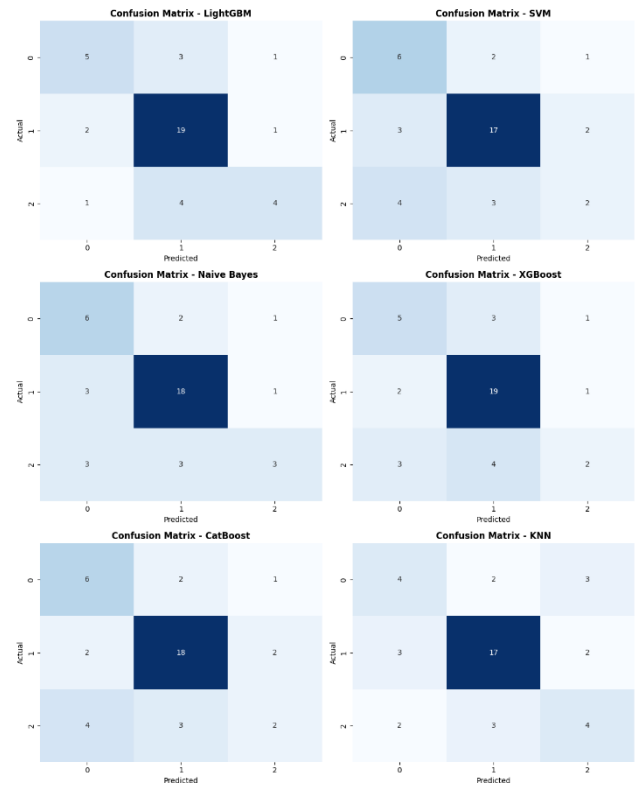
LightGBM	62%, 56%, 59%	73%, 86%, 79%	67%, 44%, 53%	70%	8,11
KNN	44%, 44%, 44%	77%, 77%, 77%	44%, 44%, 44%	62.5%	5.47
SVM	46%, 67%, 55%	77%, 77%, 77%	40%, 22%, 29%	62.5%	0.49



Gambar 4. ROC-AUC score dari setiap model

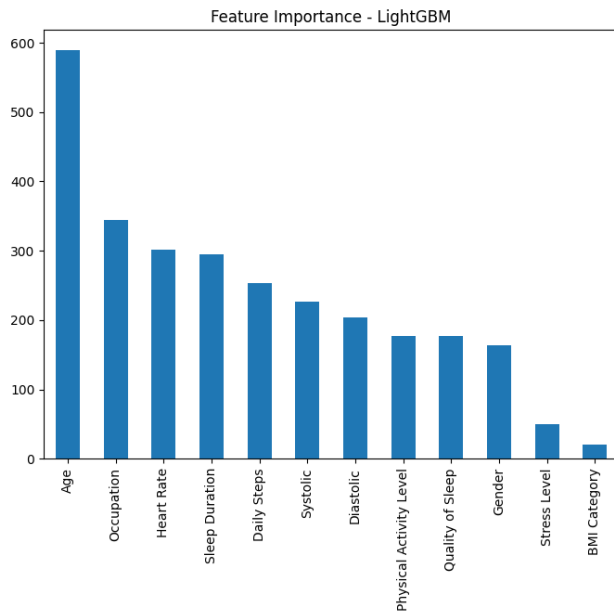
Berdasarkan hasil evaluasi, performa model pada data uji menunjukkan variasi yang cukup jelas. Algoritma KNN mencapai akurasi 62.5% dengan ROC-AUC 0.7346, SVM 62.5% dengan ROC-AUC 0.7323, Naive Bayes 67.5% dengan ROC-AUC 0.7836, XGBoost 65% dengan ROC-AUC 0.7284, CatBoost 65% dengan ROC-AUC 0.7255, dan LightGBM 70% dengan ROC-AUC 0.736. Hasil ini menunjukkan bahwa algoritma ensemble boosting, terutama LightGBM, mampu menangani kompleksitas data dengan lebih baik dibandingkan algoritma klasik seperti SVM dan KNN [15], [16].

Perbandingan training time menunjukkan Naive Bayes memiliki waktu pelatihan terpendek, 0.17 detik, sedangkan CatBoost paling lama, 21.44 detik, mencerminkan kompleksitas algoritma boosting berbasis pohon keputusan yang memerlukan iterasi pohon banyak, sementara algoritma probabilistik sederhana lebih cepat [15]. Untuk Confusion Matrix dari masing-masing model bisa dilihat pada gambar berikut.



Gambar 5. Confusion matrix model

Analisis feature importance pada LightGBM menunjukkan urutan kontribusi variabel terhadap prediksi risiko gangguan tidur sebagai berikut: Age memiliki pengaruh tertinggi, diikuti oleh Occupation, Heart Rate, Sleep Duration, Daily Steps, Systolic, Diastolic, Physical Activity Level, Quality of Sleep, Gender, Stress Level, dan BMI Category. Hasil ini menegaskan bahwa faktor usia dan pekerjaan merupakan indikator paling dominan dalam risiko gangguan tidur, diikuti oleh variabel fisiologis dan gaya hidup [16].



Gambar 6 menunjukkan kontribusi tiap fitur terhadap hasil prediksi.

Hasil uji McNemar menunjukkan statistic 1,0 dengan p-value 1,0, yang berarti perbedaan performa antar model pada dataset uji tidak signifikan secara statistik, meskipun terdapat variasi akurasi dan ROC-AUC. Temuan ini mengindikasikan bahwa seluruh model cukup stabil dalam memprediksi kelas pada data uji [36].

Secara keseluruhan, LightGBM menunjukkan performa terbaik dengan akurasi 70.00%, diikuti oleh Naïve Bayes 67.50% dan XGBoost 65.00%. Sementara itu, SVM dan KNN cenderung kurang optimal karena keterbatasan dalam menangani pola nonlinier dan sensitif terhadap distribusi data serta noise [37]. Temuan ini sejalan dengan literatur yang menunjukkan bahwa algoritma ensemble boosting unggul pada dataset kesehatan dengan variabel numerik dan kategorikal kompleks, serta mendukung interpretasi feature importance untuk analisis risiko gangguan tidur secara lebih mendalam [15], [16]. Oleh karena itu, LightGBM dan XGBoost direkomendasikan sebagai algoritma utama untuk pengembangan sistem pendukung keputusan berbasis machine learning dalam deteksi dini risiko gangguan tidur. Meskipun penelitian ini menunjukkan hasil yang menjanjikan, terdapat beberapa keterbatasan yang perlu diperhatikan. Pertama, dataset yang digunakan hanya terdiri dari 374 sampel, yang relatif kecil untuk analisis machine learning dan dapat membatasi kemampuan generalisasi temuan. Kedua, dataset mungkin tidak sepenuhnya mewakili keragaman populasi, sehingga potensi bias demografis masih ada. Ketiga, meskipun model telah divalidasi menggunakan Stratified K-Fold Cross-Validation dan Random Oversampling, pengujian lebih lanjut pada dataset yang lebih besar dan nyata diperlukan untuk memastikan ketahanan dan penerapan praktis model. Penelitian selanjutnya dapat fokus pada integrasi dataset yang lebih besar, pengukuran fisiologis

secara real-time, serta pendekatan explainable AI (XAI) untuk meningkatkan performa prediksi dan interpretabilitas dalam aplikasi klinis atau perangkat wearable.

#### IV. KESIMPULAN

Penelitian ini menunjukkan bahwa metode machine learning dapat digunakan secara efektif untuk memprediksi risiko gangguan tidur berdasarkan pola aktivitas harian dan pekerjaan. Melalui tahapan pra-pemrosesan yang meliputi pembersihan data, konstruksi data, penanganan ketidakseimbangan kelas, dan normalisasi fitur, dataset berhasil dipersiapkan dengan baik untuk proses pemodelan. Evaluasi model menggunakan metrik accuracy, precision, recall, F1-score, ROC-AUC multiclass, dan time serta analisis feature importance, memberikan gambaran komprehensif tentang performa dan faktor determinan risiko gangguan tidur.

Berdasarkan hasil evaluasi, LightGBM menunjukkan performa terbaik dengan akurasi 70,00% dan ROC-AUC 0,736, diikuti oleh Naïve Bayes dengan akurasi 67,50% dan ROC-AUC 0,784, serta XGBoost dengan akurasi 65,00% dan ROC-AUC 0,728. Analisis feature importance menunjukkan bahwa variabel Age memiliki kontribusi tertinggi terhadap prediksi, diikuti oleh Occupation, Heart Rate, Sleep Duration, Daily Steps, Systolic, Diastolic, Physical Activity Level, Quality of Sleep, Gender, Stress Level, dan BMI Category. Waktu pelatihan (training time) bervariasi, dengan Naïve Bayes paling cepat dan CatBoost paling lambat, mencerminkan kompleksitas algoritma masing-masing. Uji McNemar menunjukkan perbedaan performa antar model tidak signifikan secara statistik, menegaskan stabilitas prediksi pada dataset uji.

Berdasarkan temuan tersebut, algoritma ensemble boosting, khususnya LightGBM dan XGBoost, direkomendasikan sebagai metode yang paling efektif untuk pengembangan sistem pendukung keputusan berbasis machine learning dalam deteksi dini gangguan tidur. Temuan ini menegaskan pentingnya pemilihan algoritma yang mampu menangani data numerik dan kategorikal kompleks serta memanfaatkan analisis feature importance untuk interpretasi risiko secara mendalam.

Selain itu, implementasi model ini memiliki potensi nyata untuk aplikasi di dunia kesehatan, seperti integrasi ke wearable devices, sistem monitoring tidur di rumah sakit, atau aplikasi kesehatan digital berbasis mobile. Model prediksi ini dapat memberikan peringatan dini kepada individu dengan risiko gangguan tidur tinggi, membantu tenaga medis melakukan intervensi lebih cepat, serta mendukung pengambilan keputusan berbasis data dalam manajemen kesehatan pribadi maupun populasi.

## DAFTAR PUSTAKA

- [1] World Health Organization, "Sleep disorders: a global public health problem," *WHO*, 2021.
- [2] Sabino Cappuccio, Francesco D'Elia, Paola Strazzullo, and Michele Miller, "Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies," *Sleep*, vol. 34, no. 5, pp. 585–592, 2011.
- [3] L. M. Senaratna *et al.*, "Epidemiology of sleep disorders in adults: A systematic review," *Sleep Med.*, vol. 70, pp. 1–10, 2020.
- [4] Sanjay R. Punjabi, "The epidemiology of adult obstructive sleep apnea," *Proc. Am. Thorac. Soc.*, vol. 5, no. 2, pp. 136–143, 2008.
- [5] Yiming Li, Xinyu Wang, Qian Liu, and Jun Zhang, "Sleep apnea and cardiometabolic risk: recent epidemiological evidence," *Journal of Clinical Sleep Medicine*, vol. 18, no. 1, pp. 45–46, 2022.
- [6] Aisyah Nurlita and Sri Wijayani, "Hubungan penggunaan gadget dan pola aktivitas terhadap kualitas tidur remaja," *Jurnal Kesehatan Remaja*, vol. 3, no. 2, pp. 101–110, 2023.
- [7] Ying Zhang, Li Wei, Chen Hao, and Sun Mei, "Lifestyle factors and sleep quality among young adults: a cross-sectional study," *Int. J. Environ. Res. Public Health*, vol. 18, no. 12, p. 6301, 2021.
- [8] Sunita Mehta, "Sleep quality, cognitive function, and academic performance," *Front. Psychol.*, vol. 13, p. 1023, 2022.
- [9] Agus Putra and Rizky Hidayat, "Pengaruh pola aktivitas harian dan tekanan kerja terhadap risiko insomnia," *Jurnal Psikologi dan Kesehatan*, vol. 7, no. 1, pp. 15–24, 2024.
- [10] Soo Kim, Hyun Lee, and Min Park, "Daily activity patterns and sleep quality among working adults," *Sleep Health*, vol. 7, no. 4, pp. 345–352, 2021.
- [11] Rajesh Esteva *et al.*, "A guide to deep learning in healthcare," *Nat. Med.*, vol. 25, no. 1, pp. 24–29, 2019.
- [12] Alvin Rajkumar, Jeffrey Dean, and Ian Goodfellow, "Machine learning in medicine," *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347–1358, 2019.
- [13] Rina Maulidah and Dwi Hidayati, "Machine learning approach for sleep quality prediction based on lifestyle data," *Jurnal Informatika*, vol. 18, no. 2, pp. 55–64, 2024.
- [14] Hsiao-Chien Chien and Ming-Tsung Lee, "Predicting sleep disorders using machine learning models and lifestyle data," *Comput. Biol. Med.*, vol. 145, 2022.
- [15] X. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- [16] Guolin Ke *et al.*, "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, 2017.
- [17] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin, "CatBoost: unbiased boosting with categorical features," in *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, 2018.
- [18] Rui Liu, Kai Zhang, and Li Chen, "Ensemble boosting for health data classification: recent advances," *Comput. Biol. Med.*, vol. 134, 2021.
- [19] Thanh Tran and Minh Nguyen, "Comparative analysis of boosting algorithms on tabular health datasets," *Expert Syst. Appl.*, vol. 198, 2022.
- [20] Yifan Wang, Lin Zhao, and Xiaoming Chen, "Gradient boosting models for predicting health outcomes: a review," *J. Biomed. Inform.*, vol. 135, 2023.
- [21] Corinna Cortes and Vladimir N. Vapnik, "Support-vector networks," in *Machine Learning*, vol. 20, no. 3, 1995, pp. 273–297.
- [22] Mohammed Alshammari, Ahmed Aljohani, and Khalid Almutairi, "Almutairi Applications of classical machine learning algorithms in healthcare," *J. Healthc. Eng.*, vol. 2021, pp. 1–14, 2021.
- [23] Huy Nguyen, Anh Tran, and Linh Le, "KNN and Naive Bayes for medical data classification," *Int. J. Data Sci. Anal.*, vol. 13, pp. 45–46, 2022.
- [24] Bisong, *Google Colaboratory for Machine Learning Applications*. 2019.
- [25] M. Chen, *Hands-On Machine Learning on Google Colab*. 2022.
- [26] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. 2012.
- [27] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed. 2019.
- [28] J. Kuhn and K. Johnson, *Applied Predictive Modeling*. 2013.
- [29] P. Raschka and V. Mirjalili, *Python Machine Learning*, vol. 3rd. 2020.
- [30] Ron Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pp. 1137–1145, 1995.
- [31] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–359, 2002.
- [32] Natalia Buda, Akira Maki, and Maciej A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, 2020.
- [33] Tom Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.
- [34] L. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 2020.
- [35] A. V. Dorogush, V. Ershov, and A. Gulin, "CatBoost: gradient boosting with categorical features support," Oct. 2018.
- [36] T. Raschka, *Python Machine Learning*, 2nd ed., vol. 2018.
- [37] T. Cover and P. Hart, "Nearest Neighbor Pattern Classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967.