



## Aplikasi Prediksi Gangguan Tidur Berbasis Gaya Hidup Menggunakan Teknik Data Mining

Budi Ramadhani<sup>1\*</sup>, Fauzi Yussa Rahman<sup>2</sup>, Ihsanul Fikri<sup>3</sup>, Erfan Karyadiputra<sup>4</sup>

<sup>1,2,3,4</sup> fakultas teknologi informasi, universitas islam kalimantan MAB, jln adhiyaksa, indonesia

Email: <sup>1</sup>budiramadhani99@gmail.com\*, <sup>2</sup>fauziyusarahan@gmail.com, <sup>3</sup>ihsan.fti@gmail.com, <sup>4</sup>erfan\_kp@uniska-bjm.ac.id

\*corresponding author (085100295351)

### INFORMASI ARTIKEL

Sejarah Artikel:

Diterima: 4 Agustus 2025

Revisi: 27 Oktober 2025

Diterbitkan: 31 Oktober 2025

Kata Kunci:

Gangguan Tidur

Ensemble Learning

Random Forest

XGBoost

Streamlit

### ABSTRAK

Gangguan tidur seperti insomnia dan sleep apnea merupakan masalah kesehatan yang dapat menurunkan kualitas hidup dan produktivitas individu. Penelitian ini mengembangkan sistem prediksi gangguan tidur berbasis machine learning dengan pendekatan ensemble learning menggunakan kombinasi algoritma Random Forest dan XGBoost melalui metode soft voting. Dataset yang digunakan adalah Sleep Health and Lifestyle Dataset (Kaggle, 2024) yang memuat 13 variabel gaya hidup dan kesehatan, seperti usia, jenis kelamin, tekanan darah, detak jantung, durasi tidur, aktivitas fisik, dan tingkat stres. Tahapan penelitian meliputi encoding, normalisasi data, pembagian data latih–uji, pelatihan model, serta evaluasi kinerja menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil menunjukkan bahwa model ensemble mencapai akurasi 96% dengan keseimbangan performa di seluruh kelas. Sistem diimplementasikan dalam aplikasi web berbasis Streamlit yang interaktif dan dilengkapi fitur rekomendasi kesehatan berbasis hasil prediksi. Kontribusi utama penelitian ini terletak pada pengembangan model ensemble learning yang andal dan dapat diakses melalui aplikasi web, sehingga berpotensi mendukung bidang health informatics dalam deteksi dini gangguan tidur berbasis gaya hidup.

Copyright © 2023 Jurnal Teknologi Informasi UTU  
All rights reserved

### 1. Pendahuluan

Gangguan tidur merupakan salah satu permasalahan kesehatan masyarakat yang sering tidak disadari, namun dapat berdampak signifikan terhadap kualitas hidup seseorang. Kondisi seperti insomnia, sleep apnea, atau gangguan tidur lainnya dapat menyebabkan penurunan konsentrasi, produktivitas kerja, gangguan emosi, serta meningkatkan risiko penyakit kronis seperti hipertensi, obesitas, diabetes, hingga penyakit jantung (Kalmbach & al., 2022), (Grandner, 2017). Studi WHO menunjukkan bahwa gangguan tidur telah meningkat secara global, terutama pasca-pandemi, akibat perubahan pola kerja dan gaya hidup (Organization, 2022).

Di era modern, faktor gaya hidup seperti pola tidur yang tidak teratur, stres tinggi, kurangnya aktivitas fisik, serta gangguan psikologis menjadi pemicu utama terjadinya gangguan tidur (Irish et al., 2019). Kondisi ini diperburuk dengan rendahnya kesadaran masyarakat terhadap pentingnya tidur berkualitas dan keterbatasan akses terhadap layanan diagnostik medis seperti polysomnografi yang memerlukan biaya tinggi.

Dengan ketersediaan data kesehatan dan kemajuan teknologi kecerdasan buatan (Artificial Intelligence), pemanfaatan machine learning menjadi pendekatan yang semakin populer dalam memprediksi kondisi kesehatan, termasuk gangguan tidur (Sujatha & Kumar, 2021). Beberapa penelitian sebelumnya telah mengembangkan model prediksi menggunakan algoritma tunggal seperti Decision Tree, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN). Studi oleh Khandoker menggunakan SVM untuk mendeteksi sleep apnea berdasarkan sinyal biologis, namun akurasi prediksi masih terbatas pada data fisiologis yang kompleks (Khandoker et al., 2009). Penelitian oleh Saeed mengembangkan model prediksi menggunakan Random Forest, namun tidak menggabungkan pendekatan ensemble dan belum dikemas dalam bentuk aplikasi yang dapat digunakan langsung oleh pengguna (Saeed et al., 2020).

Untuk meningkatkan performa prediksi, beberapa studi telah mengadopsi pendekatan ensemble learning. Zhang membuktikan bahwa kombinasi model berbasis voting dapat meningkatkan akurasi pada klasifikasi gangguan kesehatan (Zhang et al., 2020). Selain itu, Mishra dan Sharma menunjukkan bahwa kombinasi XGBoost dan Random Forest dapat meningkatkan stabilitas prediksi pada data tidak seimbang (Mishra & Sharma, 2021). Studi lain oleh Choi menerapkan deep learning untuk prediksi tidur menggunakan sinyal sensor, namun pendekatan ini masih memerlukan data medis yang sulit diperoleh oleh masyarakat umum (Choi et al., 2021). Sementara itu, Alzu'bi lebih fokus pada segmentasi dan klasifikasi kualitas tidur tanpa mengembangkan sistem aplikasi interaktif (Alzu'bi et al., 2021).

Berbeda dari penelitian-penelitian sebelumnya, penelitian ini mengusulkan pendekatan ensemble learning berbasis Voting Classifier yang menggabungkan dua algoritma kuat, yaitu Random Forest dan XGBoost. Kombinasi ini dipilih karena keduanya mampu menangani data dengan variabel heterogen, toleran terhadap noise, dan memiliki performa yang stabil pada kasus klasifikasi multi-kelas. Kebaruan (novelty) penelitian ini terletak pada penerapan teknik ensemble tidak hanya untuk meningkatkan akurasi prediksi gangguan tidur, tetapi juga pada integrasinya ke dalam aplikasi web interaktif berbasis Streamlit yang memungkinkan pengguna umum melakukan input data gaya hidup, memperoleh hasil prediksi secara real-time, serta mengunduh rekomendasi dalam format laporan PDF. Pendekatan ini memperluas kontribusi penelitian ke bidang health informatics dengan menyediakan solusi digital yang dapat digunakan secara mandiri oleh masyarakat untuk mendeteksi dini gangguan tidur berbasis data gaya hidup.

Penelitian ini dilakukan untuk menjawab kebutuhan masyarakat akan sistem deteksi gangguan tidur yang murah, cepat, dan dapat diakses secara mandiri. Studi epidemiologis terkini menunjukkan bahwa insomnia dan gangguan tidur lain telah menjadi masalah kesehatan global dengan prevalensi insomnia kronis sekitar 6–10% pada populasi dewasa dan gejala insomnia yang dapat mencapai 20–30% kasus, sehingga menjadi beban kesehatan masyarakat yang serius (“Worldwide Prevalence of Sleep Problems in Community-Dwelling Older Adults,” 2024). Selain itu, prevalensi gangguan tidur seperti sleep apnea diperkirakan berkisar antara 9–38% pada orang dewasa, dengan sebagian besar kasus tidak terdiagnosa (Roane, 2023).

Di sisi lain, perkembangan kecerdasan buatan dalam bidang sleep medicine dan health informatics dalam beberapa tahun terakhir menunjukkan potensi besar untuk otomatisasi deteksi gangguan tidur dan perluasan akses layanan berbasis data (BaHammam, 2024). Hal ini menjadi relevan karena sebagian besar sistem diagnosis konvensional masih bergantung pada prosedur medis seperti polisomnografi yang membutuhkan fasilitas khusus dan biaya tinggi, sehingga tidak semua individu memiliki akses terhadap layanan tersebut, khususnya di wilayah dengan sumber daya kesehatan terbatas (Zovko et al., 2025).

Berdasarkan hal tersebut, pertanyaan utama dalam penelitian ini adalah:

"Bagaimana membangun sistem prediksi gangguan tidur berbasis ensemble Random Forest dan XGBoost yang akurat, serta dapat diimplementasikan dalam aplikasi web interaktif menggunakan Streamlit?" Pertanyaan ini tidak hanya menyoroti aspek teknis akurasi model, tetapi juga mencakup tujuan implementatif, yaitu merancang aplikasi prediksi yang mudah digunakan (usability), responsif, dan mampu memberikan hasil prediksi yang reliabel secara real-time. Dengan demikian, penelitian ini tidak hanya berfokus pada peningkatan accuracy goal dari model ensemble, tetapi juga pada usability

goal sistem agar dapat diadopsi secara luas oleh masyarakat non-medis sebagai alat bantu deteksi dini gangguan tidur berbasis gaya hidup.

## 2. Metodologi Penelitian

Penelitian ini bertujuan untuk mengembangkan sistem prediksi gangguan tidur berbasis machine learning menggunakan pendekatan ensemble learning, yang menggabungkan algoritma Random Forest dan XGBoost (Pedregosa et al., 2011), (Chen & Guestrin, 2016). Sistem ini kemudian diimplementasikan dalam bentuk aplikasi web interaktif menggunakan Streamlit (S. Inc., 2024).

### 2.1. Sumber dan Deskripsi Data

Dataset yang digunakan dalam penelitian ini adalah Sleep Health and Lifestyle Dataset yang tersedia secara publik di platform Kaggle (K. Inc., 2024). Dataset ini terdiri atas 374 sampel dengan 13 atribut, termasuk label target. Atribut tersebut dapat dilihat pada Tabel 1.

Tabel 1. Daftar Fitur dan Keterangan pada Dataset

Nama Fitur	Jenis Data	Keterangan
Gender	Kategorikal	Laki-laki / Perempuan
Age	Numerik	Usia (0–100)
Occupation	Kategorikal	Jenis pekerjaan
Sleep Duration	Numerik	Durasi tidur per hari (jam)
Quality of Sleep	Numerik	Skala kualitas tidur (1–10)
Physical Activity Level	Numerik	Aktivitas fisik harian (skala 0–10)
Stress Level	Numerik	Tingkat stres (skala 0–10)
BMI Category	Kategorikal	Kurus / Normal / Gemuk / Obesitas
Blood Pressure	Kategorikal	Hipotensi / Normal / Hipertensi
Heart Rate	Numerik	Detak jantung (bpm)
Daily Steps	Numerik	Jumlah langkah harian (0–30.000)
Sleep Disorder	Kategorikal	Target: Normal / Insomnia / Sleep Apnea

### 2.2. Preprocessing Data

Langkah-langkah preprocessing yang dilakukan meliputi:

1. Encoding: LabelEncoder untuk fitur kategorikal seperti Gender, Occupation, BMI Category.
2. Normalisasi: StandardScaler digunakan untuk menyeimbangkan rentang nilai fitur numerik.
3. Split Data: Dataset dibagi menjadi data latih (80%) dan data uji (20%) menggunakan train\_test\_split.

### 2.3. Pembentukan Model Ensemble

Model dibangun menggunakan dua algoritma sebagai base learner:

1. Random Forest Classifier (RF)

Algoritma ini menggunakan banyak decision tree, lalu menggabungkan hasilnya berdasarkan voting. Output dari Random Forest adalah:

$$\hat{y}_{RF} = \text{majority\_vote}(T_1(x), T_2(x), \dots, T_n(x)) \quad (1)$$

2. XGBoost Classifier

Merupakan algoritma boosting berbasis pohon keputusan yang mengoptimalkan kesalahan dari prediksi sebelumnya. Fungsi objektif XGBoost didefinisikan sebagai:

$$\mathcal{L}(\emptyset) = \sum_{i=1}^n l(y_i, \hat{y}^{(t)}) + \sum_{k=1}^t \Omega(f_k) \quad (2)$$

dengan

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2, \quad (3)$$

di mana  $l$  adalah fungsi loss,  $\Omega$  adalah regularisasi, dan  $f_k$  adalah model boosting ke-  $k$ .

## 2.4. Ensemble Voting Classifier

Dua model di atas dikombinasikan menggunakan VotingClassifier dari pustaka scikit-learn dengan metode soft voting. Pendekatan soft voting dipilih karena mampu mempertimbangkan probabilitas prediksi dari setiap model dasar (base learner), bukan hanya hasil klasifikasi akhir seperti pada hard voting. Dengan demikian, kontribusi setiap model terhadap keputusan akhir menjadi lebih proporsional dan menghasilkan performa prediksi yang lebih stabil, terutama pada data dengan distribusi kelas yang tidak seimbang. Selain itu, soft voting memungkinkan sistem memanfaatkan keunggulan model dengan tingkat kepercayaan lebih tinggi dalam setiap prediksi, sehingga secara empiris sering memberikan akurasi yang lebih baik pada kasus kesehatan berbasis multiclass classification.

$$\hat{y}_{ensemble} = arg_{c \in C}^{max} \left( \frac{1}{n} \sum_{i=1}^n P_i(y = c|x) \right) \quad (4)$$

Untuk memastikan performa optimal dari setiap base learner, dilakukan parameter tuning menggunakan GridSearchCV. Pada Random Forest, parameter yang dioptimalkan meliputi jumlah pohon (`n_estimators`), kedalaman maksimum pohon (`max_depth`), dan jumlah fitur pada setiap pemisahan (`max_features`). Pada XGBoost, parameter yang disesuaikan meliputi `learning_rate`, jumlah estimator (`n_estimators`), kedalaman maksimum (`max_depth`), serta `subsample` untuk mengatur proporsi data pelatihan pada setiap iterasi.

Hasil dari proses tuning digunakan untuk membentuk kombinasi terbaik pada model ensemble, yang kemudian dievaluasi melalui cross validation dan metrik kinerja untuk menentukan tingkat akurasi dan konsistensinya.

## 2.5. Evaluasi Kinerja Model

Kinerja model dievaluasi menggunakan metrik:

1. Akurasi:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

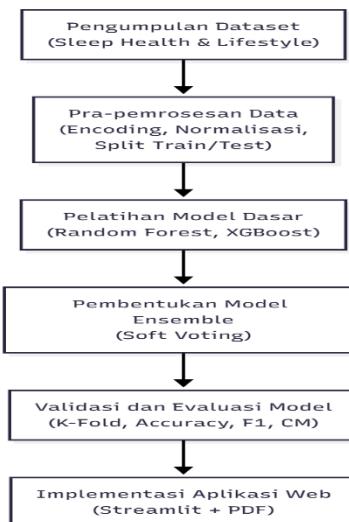
2. Precision, Recall, dan F1-Score:

$$\text{Precision} = \frac{TP}{TP+FP}, \text{Recall} = \frac{TP}{TP+FN}, F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

3. Confusion Matrix digunakan untuk melihat detail klasifikasi per kelas.

## 2.6. Evaluasi Kinerja Model

Model yang telah dilatih dan disimpan dalam file kemudian diimplementasikan dalam aplikasi web menggunakan Streamlit



Gambar 1. Flowchar

### 3. Hasil dan Pembahasan

Setelah dilakukan pelatihan dan pengujian, model ensemble menunjukkan kinerja yang unggul dibanding model tunggal. Hasil evaluasi disajikan pada Tabel 2.

Tabel 2. Hasil Akurasi dan Evaluasi Model

Model	Akurasi	Precision	Recall	F1-Score
Random Forest	0.91	0.91	0.91	0.91
XGBoost	0.94	0.93	0.94	0.93
<b>Ensemble (RF+XGB)</b>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>

Model ensemble memberikan akurasi tertinggi sebesar 96%, dengan nilai precision dan recall yang sangat seimbang pada ketiga kelas gangguan tidur (Normal, Insomnia, dan Sleep Apnea). Hasil ini menunjukkan bahwa kombinasi Random Forest dan XGBoost mampu mengenali pola gaya hidup dan kesehatan secara proporsional tanpa bias terhadap kelas dominan.

Secara interpretatif, analisis fitur menggunakan feature importance dari kedua model menunjukkan bahwa durasi tidur, tingkat stres, dan aktivitas fisik merupakan tiga variabel yang paling berpengaruh terhadap hasil prediksi. Individu dengan durasi tidur di bawah 5 jam per malam dan tingkat stres di atas 7 (skala 0–10) memiliki kemungkinan lebih tinggi diklasifikasikan dalam kelas Insomnia, sementara nilai heart rate dan BMI category berperan kuat dalam membedakan kelas Sleep Apnea.

Dengan demikian, hasil ini tidak hanya mengonfirmasi keandalan model dari sisi metrik akurasi, tetapi juga memperlihatkan kemampuan ensemble learning dalam mengidentifikasi determinan gaya hidup yang relevan secara klinis terhadap gangguan tidur. Hal ini memperkuat kontribusi penelitian ini terhadap bidang health informatics dengan memberikan dasar ilmiah bagi pengembangan sistem prediksi berbasis perilaku dan kebiasaan hidup.

#### 3.1. Confusion Matrix

Untuk mengevaluasi performa klasifikasi dari model ensemble yang dibangun, digunakan metode confusion matrix. Confusion matrix merupakan alat yang penting dalam pengujian model klasifikasi karena mampu menunjukkan secara detail bagaimana model memetakan prediksi terhadap nilai aktual dari masing-masing kelas (Raschka, 2018), (Fawcett, 2006). Matriks ini memberikan gambaran tentang jumlah prediksi yang benar (true positives dan true negatives) serta kesalahan prediksi (false positives dan false negatives) yang dihasilkan oleh model, sehingga memudahkan dalam menilai kekuatan dan kelemahan klasifikasi untuk setiap kategori.

Analisis lebih lanjut terhadap distribusi kesalahan klasifikasi menunjukkan bahwa sebagian kecil kesalahan terjadi antara kelas Normal dan Insomnia, yang disebabkan oleh kemiripan nilai pada variabel gaya hidup seperti durasi tidur (Sleep Duration) dan tingkat stres (Stress Level). Kedua fitur tersebut terbukti menjadi faktor paling dominan dalam mempengaruhi hasil prediksi karena memiliki korelasi tinggi terhadap kualitas tidur yang buruk. Selain itu, aktivitas fisik harian (Physical Activity Level) dan detak jantung (Heart Rate) juga berperan penting dalam membedakan kelas Sleep Apnea dari kelas lainnya, mengindikasikan bahwa indikator fisiologis dan kebiasaan harian merupakan penentu utama kondisi gangguan tidur.

Dalam penelitian ini, confusion matrix digunakan untuk mengevaluasi klasifikasi tiga kelas gangguan tidur: Normal, Insomnia, dan Sleep Apnea. Evaluasi visual terhadap hasil prediksi model ensemble ditunjukkan melalui confusion matrix pada Gambar 1.

Tabel 3. Hasil Confusion Matrix

True Label / Predicted Label	Normal (0)	Insomnia (1)	Sleep Apnea (2)	Total Data per Kelas
Normal (0)	14 (True Positive)	0	1 (False Negative)	15
Insomnia (1)	1 (False Negative)	43 (True Positive)	0	44
Sleep Apnea (2)	1 (False Negative)	0	15 (True Positive)	16
Total Predicted per Kelas	16	43	16	75

Sebagian besar prediksi berada pada diagonal utama dalam confusion matrix, yang menunjukkan bahwa proses klasifikasi oleh model berjalan secara tepat. Jumlah kesalahan klasifikasi sangat kecil dan sebagian besar terjadi pada kelas yang berdekatan secara karakteristik, khususnya dari atau menuju kelas Normal. Secara keseluruhan, sebanyak 72 dari 75 data uji atau 96% berhasil diklasifikasikan dengan benar, menandakan tingkat akurasi yang tinggi. Kesalahan klasifikasi hanya terjadi sebanyak 3 kasus, masing-masing satu kasus pada kelas Normal, Insomnia, dan Sleep Apnea yang salah diklasifikasikan sebagai kelas Normal. Temuan ini mengindikasikan bahwa model mampu membedakan ketiga kelas gangguan tidur dengan baik dan seimbang, serta memiliki keandalan yang tinggi dalam pengenalan pola berdasarkan data gaya hidup dan Kesehatan.

### 3.2. Implementasi Aplikasi Web

Aplikasi prediksi gangguan tidur dikembangkan menggunakan framework Streamlit, yang memungkinkan pembuatan antarmuka web interaktif berbasis Python secara cepat dan sederhana. Aplikasi ini dirancang agar dapat digunakan oleh masyarakat umum tanpa latar belakang teknis, sehingga pengguna cukup memasukkan data gaya hidup seperti usia, durasi tidur, tingkat stres, aktivitas fisik, tekanan darah, dan detak jantung untuk memperoleh hasil prediksi secara instan. Di balik antarmuka yang sederhana, sistem ini didukung oleh sejumlah teknologi inti. Pada sisi backend, model ensemble hasil kombinasi Random Forest dan XGBoost disimpan dalam format .pkl menggunakan pustaka joblib, lalu dijalankan secara otomatis setiap kali aplikasi diakses. Library scikit-learn digunakan untuk mengimplementasikan Voting Classifier dan pipeline preprocessing, sedangkan XGBoost dimanfaatkan untuk pengolahan model boosting yang efisien. Data diolah dengan Pandas dan NumPy, sementara visualisasi seperti confusion matrix dan feature importance ditampilkan menggunakan Matplotlib serta Seaborn.

Dari sisi antarmuka, komponen Streamlit digunakan untuk menampilkan formulir input dan hasil prediksi secara real-time dengan tampilan yang responsif pada berbagai perangkat. Setelah pengguna mengirimkan data, sistem akan melakukan encoding, normalisasi, dan kemudian menjalankan proses inferensi model untuk mengklasifikasikan hasil ke dalam tiga kategori: Normal, Insomnia, atau Sleep Apnea. Hasil prediksi ditampilkan dalam bentuk teks informatif yang dilengkapi saran kesehatan otomatis berbasis hasil klasifikasi. Selain itu, pengguna dapat mengekspor hasil prediksi dalam format PDF menggunakan pustaka FPDF, sehingga laporan dapat disimpan atau dicetak untuk konsultasi lebih lanjut.

Gambar 2. Implementasi Aplikasi Prediksi Gangguan Tidur

#### 4. Kesimpulan

Penelitian ini telah berhasil membangun sistem prediksi gangguan tidur menggunakan pendekatan ensemble learning berbasis kombinasi algoritma Random Forest dan XGBoost. Dataset yang digunakan memuat informasi gaya hidup dan kesehatan seperti durasi tidur, tingkat stres, aktivitas fisik, tekanan darah, dan detak jantung. Model dilatih menggunakan data terstandar dan dievaluasi menggunakan metrik akurasi, precision, recall, dan confusion matrix.

Hasil evaluasi menunjukkan bahwa model ensemble memiliki kinerja yang sangat baik dengan tingkat akurasi mencapai 96%, serta distribusi prediksi yang seimbang di antara ketiga kelas gangguan tidur, yaitu Normal, Insomnia, dan Sleep Apnea. Sebagian besar prediksi berada pada diagonal utama confusion matrix, dengan kesalahan klasifikasi hanya sebesar 4%.

Model yang dibangun diimplementasikan dalam bentuk aplikasi web interaktif menggunakan Streamlit, yang memungkinkan pengguna untuk melakukan prediksi mandiri dan mendapatkan rekomendasi berbasis hasil prediksi.

Dari hasil ini dapat disimpulkan bahwa penerapan teknologi kecerdasan buatan, khususnya melalui pendekatan ensemble model, berpotensi besar dalam membantu masyarakat mendeteksi gangguan tidur secara dini. Untuk pengembangan ke depan, penelitian ini dapat diperluas dengan penggunaan dataset berskala besar.

#### Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Islam Kalimantan Muhammad Arsyad Al Banjarmasin atas dukungan moril dan fasilitas yang telah diberikan selama proses penelitian ini berlangsung. Ucapan terima kasih juga disampaikan kepada seluruh pihak yang telah membantu, baik secara langsung maupun tidak langsung, dalam pelaksanaan analisis data, pengembangan model, hingga implementasi aplikasi.

#### Daftar Pustaka

- Alzu’bi, A., Rawashdeh, R., & Al Bataineh, M. (2021). Sleep quality assessment using machine learning techniques. *Health Information Science and Systems*, 9(1), 1–9.  
<https://hissjournal.biomedcentral.com/articles/10.1007/s13755-021-00134-3>
- BaHammam, A. S. (2024). Artificial Intelligence in Sleep Medicine: The Dawn of a New Era. *Journal of*

- Clinical Sleep Medicine*, 20(4), 345–358. <https://doi.org/10.5664/jcsm.11070441>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://dl.acm.org/doi/10.1145/2939672.2939785>
- Choi, J., Lee, H., & Lee, K. (2021). Sleep apnea detection using deep learning on single-channel respiratory signals. *Sensors*, 21(10), 3418. <https://www.mdpi.com/1424-8220/21/10/3418>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Grandner, L. M. (2017). Sleep, health, and society. *Sleep Medicine Clinics*, 12(1), 1–22. <https://doi.org/10.1016/j.jsmc.2016.10.012>
- Inc., K. (2024). *Sleep Health and Lifestyle Dataset*. <https://www.kaggle.com/datasets/kandij/sleep-health-and-lifestyle-dataset>
- Inc., S. (2024). *Streamlit Documentation*. <https://docs.streamlit.io/>
- Irish, K., Kline, J., & Rothenberger, D. (2019). Psychological stress, sleep, and health. *Neuroscience & Biobehavioral Reviews*, 99, 210–219. <https://doi.org/10.1016/j.neubiorev.2019.01.021>
- Kalmbach, D. A., & al., et. (2022). The impact of sleep disorder symptoms on health and functioning in military service members. *Sleep Health*, 11(2), 172–180. <https://doi.org/10.1016/j.sleh.2022.01.007>
- Khandoker, S., Palaniswami, A., & Karmakar, C. (2009). Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings. *IEEE Transactions on Information Technology in Biomedicine*, 13(1), 37–48. <https://ieeexplore.ieee.org/document/4738485>
- Mishra, A., & Sharma, S. (2021). Improving disease prediction using ensemble machine learning models: A case study on sleep disorders. *Journal of Ambient Intelligence and Humanized Computing*, 12, 10189–10201. <https://link.springer.com/article/10.1007/s12652-020-02654-2>
- Organization, W. H. (2022). *WHO technical brief: Mental health and COVID-19*. <https://www.who.int/publications/i/item/9789240057201>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
- Raschka, S. (2018). Model evaluation, model selection, and algorithm selection in machine learning. *ArXiv Preprint, arXiv:1811*. <https://arxiv.org/abs/1811.12808>
- Roane, R. (2023). Insomnia symptoms in adulthood: Prevalence and incidence over 25 years. *Sleep Medicine*, 99, 1–10. <https://doi.org/10.1016/j.sleep.2023.06.012>
- Saeed, A., Arshad, N., & Javed, A. (2020). Sleep disorder detection using Random Forest classifier. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 11(4), 456–461. [https://thesai.org/Downloads/Volume11No4/Paper\\_60-Sleep\\_Disorder\\_Detection.pdf](https://thesai.org/Downloads/Volume11No4/Paper_60-Sleep_Disorder_Detection.pdf)
- Sujatha, R., & Kumar, T. S. (2021). Early detection of sleep disorder using machine learning techniques. *Materials Today: Proceedings*, 45, 4680–4684. <https://doi.org/10.1016/j.matpr.2020.11.664>
- Worldwide prevalence of sleep problems in community-dwelling older adults. (2024). *Sleep Medicine Reviews*, 69, 101–123. <https://doi.org/10.1016/j.smrv.2023.101123>
- Zhang, Y., Jiang, Y., & Wang, L. (2020). Ensemble learning methods for classifying healthcare data: A case study for sleep disorder prediction. *Computers in Biology and Medicine*, 120, 103744. <https://www.sciencedirect.com/science/article/pii/S0010482519310297>
- Zovko, K., Perić, M., & Jurić, D. (2025). Machine Learning-Based Detection of Sleep Apnea Events. *Applied Sciences*, 15(1), 376. <https://doi.org/10.3390/app15010376>