Sistem Deteksi *Sleep Disorder* Berbasis Random Forest

1st Amanda Oktaviana Saragih
Informatika
Universitas Telkom Surabaya
Surabaya, Indonesia
saragiha@student.telkomuniversity.ac.id

2nd Kharisma Monika Dian Pertiwi
Informatika
Universita Telkom Surabaya
Surabaya, Indonesia
kharismamonikadp@telkomuniversity.ac.id

3rd Daud Muhajir
Informatika
Universitas Telkom Surabaya
Surabaya, Indonesia
daudmuhajir@telkomuniversity.ac.id

***Abstrak* — sleep disorder merupakan gangguan tidur yang dapat berdampak signifikan terhadap kesehatan fisik maupun mental seseorang. Banyak individu mengalami gangguan tidur tanpa menyadarinya, sementara pemeriksaan klinis memerlukan biaya dan waktu yang relatif besar. Oleh karena itu, diperlukan sistem yang mampu mendukung proses deteksi sleep disorder secara praktis. Penelitian ini bertujuan untuk merancang sistem deteksi sleep disorder menggunakan algoritma Random Forest. Data yang digunakan berasal dari dataset Sleep Health and Lifestyle dari Kaggle dengan tiga kelas, yaitu None, Insomnia, dan Sleep Apnea. Fitur yang digunakan meliputi durasi tidur, kualitas tidur, tingkat stres, aktivitas fisik, kategori BMI, detak jantung, dan jumlah langkah harian. Model dikembangkan menggunakan supervised learning dengan pembagian data train dan test. Hasil pengujian menunjukkan bahwa Random Forest menghasilkan performa yang cukup baik dengan akurasi sebesar 0,7586, precision 0,77, recall 0,78, dan F1-score 0,76. Model kemudian diimplementasikan ke dalam aplikasi berbasis web menggunakan Flask sehingga pengguna dapat melakukan deteksi kondisi tidurnya secara mandiri. Sistem ini diharapkan dapat membantu masyarakat dalam memantau kualitas tidur secara lebih mudah dan terjangkau.**

***Kata kunci*— sleep disorder, random forest, machine learning, klasifikasi, web.**

I. PENDAHULUAN

Sleep disorder merupakan permasalahan kesehatan yang ditandai dengan terganggunya pola tidur yang dapat berdampak pada kualitas hidup, konsentrasi, serta kondisi fisik dan mental. Jenis gangguan tidur seperti insomnia dan sleep apnea sering kali tidak disadari oleh penderitanya sehingga tidak terdeteksi sejak dini. Di sisi lain, pemeriksaan klinis memerlukan biaya dan waktu yang cukup besar sehingga belum praktis bagi masyarakat.

Perkembangan teknologi wearable memungkinkan pemantauan kondisi tidur secara non-invasif melalui pencatatan parameter fisiologis seperti durasi tidur, detak jantung, kualitas tidur, serta aktivitas fisik. Data tersebut dapat dianalisis menggunakan teknik machine learning untuk mengenali pola gangguan tidur.

Beberapa penelitian menunjukkan bahwa algoritma Random Forest memiliki performa yang baik dalam klasifikasi gangguan tidur, khususnya pada dataset yang memiliki kombinasi fitur numerik dan kategorikal. Namun, sebagian penelitian masih berfokus pada evaluasi model tanpa implementasi sistem yang dapat digunakan langsung oleh pengguna.

Oleh karena itu, penelitian ini berfokus pada pengembangan sistem deteksi sleep disorder berbasis Random Forest yang diimplementasikan dalam aplikasi web. Sistem ini diharapkan dapat membantu pengguna memperoleh informasi awal mengenai kondisi tidurnya secara mandiri.

II. KAJIAN TEORI

A. Sleep Disorder

Sleep disorder merupakan gangguan pada pola tidur yang dapat memengaruhi kesehatan fisik dan mental, seperti insomnia dan sleep apnea. Gangguan ini berhubungan dengan kualitas tidur, durasi tidur, serta kondisi fisiologis seseorang.

B. Random Forest

Random Forest adalah algoritma ensemble learning yang mengombinasikan banyak pohon keputusan dengan teknik bootstrap sampling dan pemilihan fitur acak. Hasil prediksi ditentukan melalui majority voting sehingga mampu mengurangi overfitting dan meningkatkan kestabilan model.

C. Data Sleep Health and Lifestyle

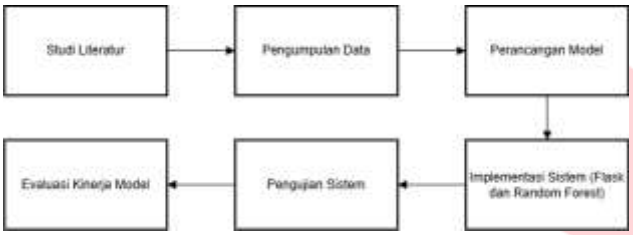
Dataset ini berisi data kesehatan dan gaya hidup individu, meliputi gender, usia, durasi tidur, kualitas tidur, aktivitas fisik, stres, kategori BMI, detak jantung, jumlah langkah harian, serta label sleep disorder.

D. Flask

Flask merupakan framework Python yang digunakan untuk membangun aplikasi web ringan dan mudah diintegrasikan dengan model machine learning sehingga cocok digunakan untuk implementasi sistem klasifikasi.

III. METODE

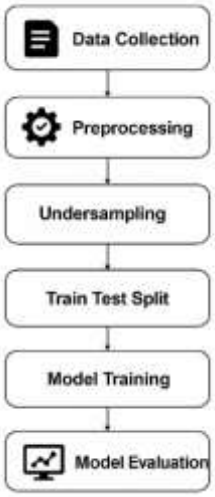
Penelitian ini menggunakan dataset Sleep Health and Lifestyle dari Kaggle sebagai sumber data utama. Tahapan penelitian meliputi preprocessing data, seleksi fitur, pembagian data, pelatihan model, evaluasi, serta implementasi sistem. Alur penelitian secara umum ditunjukkan pada Gambar 1.



GAMBAR 1
(ALUR PENELITIAN)

Preprocessing meliputi pembersihan data, encoding fitur kategorikal, normalisasi fitur numerik, dan penyelarasan struktur data. Fitur yang digunakan adalah gender, age, sleep duration, quality of sleep, physical activity level, stress level, BMI category, heart rate, dan daily steps.

Untuk mengatasi ketidakseimbangan kelas digunakan metode Random UnderSampling. Data kemudian dibagi menjadi 75% data latih dan 25% data uji. Model dilatih menggunakan Random Forest dengan parameter n_estimators = 150 dan criterion = gini. Proses pelatihan model random forest ditunjukkan pada Gambar 2.



GAMBAR 2
(MODEL RANDOM FOREST)

Evaluasi dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Model yang telah dilatih kemudian diintegrasikan ke dalam aplikasi web berbasis Flask sehingga pengguna dapat memasukkan data dan memperoleh hasil prediksi sleep disorder. Arsitektur system aplikasi ditunjukkan pada Gambar 3.



GAMBAR 3
(ARSITEKTUR SYSTEM APLIKASI)

IV. HASIL DAN PEMBAHASAN

Hasil pengujian menunjukkan bahwa model Random Forest mampu mengklasifikasikan sleep disorder dengan performa yang cukup baik. Nilai akurasi yang diperoleh sebesar 0,7586 dengan precision 0,77, recall 0,78, dan F1-score 0,76. Hasil evaluasi model ditunjukkan pada Tabel 1.

Metrik	Nilai
Accuracy	0.7586
Precision	0.77
Recall	0.78
F1-Score	0.76

TABEL 1
(Hasil Evaluasi Model)

Model menunjukkan kestabilan dalam membedakan tiga kelas, yaitu None, Insomnia, dan Sleep Apnea. Kesalahan klasifikasi masih terjadi pada beberapa data yang memiliki karakteristik mirip antar kelas, khususnya antara insomnia dan none.

Implementasi ke dalam sistem web memungkinkan pengguna melakukan input data kesehatan secara mandiri. Sistem kemudian memproses data melalui preprocessing dan model Random Forest untuk menghasilkan prediksi yang ditampilkan pada dashboard. Tampilan hasil prediksi pada aplikasi web ditunjukkan pada Gambar 4.



GAMBAR 4
(HASIL PREDIKSI)

V. KESIMPULAN

Penelitian ini berhasil membangun sistem deteksi sleep disorder berbasis Random Forest menggunakan dataset Sleep Health and Lifestyle. Model menghasilkan performa yang cukup baik dengan akurasi 75,86% serta nilai precision, recall, dan F1-score yang stabil. Implementasi ke dalam aplikasi web berbasis Flask memungkinkan pengguna melakukan deteksi kondisi tidur secara mandiri. Sistem ini dapat dimanfaatkan sebagai alat bantu pemantauan awal gangguan tidur dan masih dapat dikembangkan lebih lanjut dengan integrasi data wearable dan optimasi model.

REFERENSI

[1] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[2] E. Elgammal, “Sleep Health and Lifestyle Dataset,” Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/ealaxi/paysim1 [Accessed: Jun. 2024].

[3] S. Raschka and V. Mirjalili, Python Machine Learning, 3rd ed. Birmingham, UK: Packt Publishing, 2019.

[4] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd ed. Sebastopol, CA: O’Reilly Media, 2019.

[5] M. Grinberg, Flask Web Development: Developing Web Applications with Python, 2nd ed. Sebastopol, CA: O’Reilly Media, 2018.

[6] H. Kaur and V. Singh, “Detection of sleep disorders using machine learning techniques,” International Journal of Computer Applications, vol. 176, no. 34, pp. 1–6, 2020.

[7] A. K. Jain, J. Mao, and K. M. Mohiuddin, “Artificial neural networks: A tutorial,” Computer, vol. 29, no. 3, pp. 31–44, 1996.

[8] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[9] S. Redline and T. Young, “Epidemiology and natural history of sleep-disordered breathing,” Sleep Medicine Reviews, vol. 7, no. 5, pp. 419–433, 2003.

[10] I. H. Witten, E. Frank, and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 3rd ed. Burlington, MA: Morgan Kaufmann, 2011.