

Prediksi *Sleep Disorder* Menggunakan Metode Algoritma *Naïve Bayes Gaussian*

Galuh Dwi Rahayuningtyas

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Gresik,
Jawa Timur, Indonesia.

Email : gd.rhy47@gmail.com

Abstract. *Sleep disorders are conditions characterized by disturbances in the quantity, quality, and/or duration of sleep. Common sleep disorders are insomnia and sleep apnea. If left untreated, insomnia and sleep apnea can become chronic and lead to other more serious disease symptoms such as hypertension, depression, physical decline and also vital functions that can cause other disease complications. With the existence of a system that can help predict the diagnosis of sleep disorders will help find out the sleep disorders suffered and can be treated immediately. The prediction process is carried out using the Naïve Bayes Gaussian algorithm classification method. Based on testing training data and test data, an accuracy of 85.3% is obtained. This proves that classification using the Naïve Bayes Gaussian algorithm is good enough to predict sleep disorders.*

Keywords : *Sleep disorders, Insomnia, Sleep Apnea, Naïve Bayes Gaussian, Classification*

Abstrak. Gangguan pada tidur adalah suatu kondisi yang ditandai dengan adanya gangguan pada kuantitas, kualitas, dan atau durasi tidur. Gangguan tidur yang umum terjadi adalah insomnia dan sleep apnea. Jika tidak ditangani, insomnia maupun sleep apnea dapat menjadi kronis dan mengakibatkan terkena gejala penyakit yang lebih serius lainnya seperti hipertensi, depresi, penurunan fisik dan juga fungsi vital yang dapat menyebabkan komplikasi penyakit lainnya. Dengan adanya sebuah sistem yang dapat membantu prediksi diagnosis gangguan tidur akan membantu mengetahui gangguan tidur yang diderita dan dapat ditangani dengan segera. Proses prediksi dilakukan menggunakan metode klasifikasi algoritma Naïve Bayes Gaussian. Berdasarkan pengujian data training dan data uji didapatkan akurasi sebesar 85,3%. Hal tersebut membuktikan bahwa klasifikasi menggunakan algoritma Naïve Bayes Gaussian tergolong baik untuk memprediksi gangguan tidur.

Kata kunci : Gangguan tidur, Insomnia, Sleep Apnea, Naïve Bayes Gaussian, Klasifikasi

1. PENDAHULUAN

Gangguan pada tidur adalah suatu kondisi yang ditandai dengan adanya gangguan pada kuantitas, kualitas, dan atau durasi tidur. Gangguan tidur yang umum terjadi adalah insomnia dan sleep apnea. . Insomnia ditandai dengan kesulitan untuk memulai atau mempertahankan tidur (Bollu & Kaur, 2019). Sleep apnea adalah gangguan tidur dimana ditandai dengan kesulitan bernapas saat tidur. Kondisi ini juga menimbulkan beberapa gejala lainnya seperti rasa tercekik dan nafas terengah-engah saat tidur, penurunan konsentrasi, dan gejala merugikan lainnya (Azzahra, 2019). Jika tidak ditangani, insomnia maupun sleep apnea dapat mempengaruhi aktifitas dan kesehatan fisik. Gangguan tidur seperti insomnia dan sleep apnea memerlukan diagnosis dengan melihat faktor-faktor yang dapat mempengaruhi agar dapat ditangani dengan tepat.

Kendala ekonomi ataupun jarak menjadikan masyarakat enggan pergi untuk periksa dan mendapatkan diagnosis yang tepat. Untuk memudahkan masyarakat mendapatkan diagnosis yang tepat dengan mudah diperlukan sebuah sistem yang dapat membantu prediksi diagnosis gangguan tidur. Data mining adalah suatu proses pengumpulan dan pengolahan data yang efektif untuk digunakan sebagai metode dalam memprediksi gangguan tidur. Salah satu algoritma yang sangat sering digunakan dan menghasilkan performa baik adalah algoritma naive bayes. Naïve bayes adalah teknik klasifikasi yang didasarkan pada penerapan teorema Bayes. Algoritma ini merupakan metode klasifikasi yang cukup mudah dan akurat untuk digunakan dalam proses klasifikasi (Mubarog, et al., 2019). Naïve bayes sering digunakan dalam sistem klasifikasi untuk memprediksi penyakit. Dalam penelitian terdahulu oleh Karim dkk yang membuat sistem pakar untuk mengidentifikasi gejala stroke, naïve bayes memiliki performa yang sangat baik (Karim, et al., 2021). Penelitian sebelumnya oleh Diana Sari untuk memprediksi gangguan tidur pada dataset Sleep Health and Lifestyle dengan menggunakan Support Vector Machine (SVM) dan Neural Network didapatkan hasil akurasi yang tinggi 90.1% pada SVM dan 91.2% pada nilai akurasi Neural Network. Penerapan algoritma lainnya untuk memprediksi penyakit gangguan tidur sangat diperlukan untuk melihat bagaimana hasil dari algoritma lain, yang dalam penelitian ini adalah naive bayes.

2. TINJAUAN PUSTAKA

Data Mining

Data mining adalah suatu proses pengumpulan dan pengolahan data menggunakan teknik atau metode tertentu yang bertujuan mengekstraksi atau mengambil suatu informasi dari data tersebut dengan memperhatikan pola dan hubungan antar data dalam jumlah besar (Sinaga, et al., 2021). Menurut Charles Zai (2022), data mining merupakan sebuah proses perhitungan semi otomatis yang digunakan untuk mengekstraksi dan menggali informasi dalam suatu data besar dengan menggunakan teknik statistik, kecerdasan buatan, dan machine learning.

Klasifikasi

Menurut Putri dan Wijayanto (2022), klasifikasi merupakan proses untuk menemukan fungsi yang dapat menjelaskan dan membedakan kelas data. Klasifikasi adalah proses evaluasi data yang bertujuan untuk mengkategorikan data ke dalam salah satu dari beberapa kelas yang telah ditentukan. Tujuan dari metode klasifikasi adalah memahami sebuah dataset sehingga kita dapat membuat aturan atau model yang mampu mengklasifikasikan data baru yang belum pernah kita lihat sebelumnya.

Naïve Bayes

Naïve Bayes merupakan suatu metode pengklasifikasian dalam data mining dengan menggunakan probabilitas sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dataset yang diberikan (Alkhairi et al., 2021). Teorema Bayes dirumuskan dengan formula sebagai berikut :

$$P(X | H) = \frac{P(X | H) \cdot P(H)}{P(X)}$$

Dimana :

X : Data dengan class yang belum diketahui
H : Hipotesis data X

$P(H | X)$: Probabilitas hipotesis H berdasarkan kondisi X

$P(X | H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(H)$: Probabilitas hipotesis H

$P(X)$: Probabilitas X

Persamaan diatas merupakan model teorema naïve bayes yang selanjutnya digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinu digunakan rumus Densitas Gauss sebagai berikut:

$$P(X_i = x_i | Y = y_i) = \frac{1}{2\pi\sigma_{ij}^2} e\left(-\frac{(x_i - \mu_i)^2}{2\pi\sigma_{ij}^2\sigma\sigma}\right)$$

Keterangan:

P : Peluang

Xi : atribut ke i

xi : nilai atribut ke i

Y : kelas yang dicari

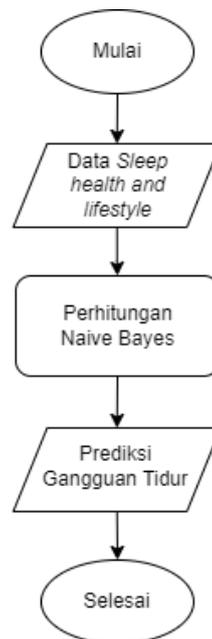
yi : subkelas Y yang dicari

π : mean, rata-rata dari seluruh atribut

σ : standar deviasi, menyatakan variasi dari seluruh atribut.

3. METODE PENELITIAN

Dalam perhitungan klasifikasi pada sistem ini akan menggunakan algoritma Naïve Bayes. Berikut adalah alur dari sistem prediksi gangguan tidur yang akan dibuat menggunakan algoritma Naïve Bayes.



Sistem ini memiliki cara kerja yang sederhana. Pengguna hanya perlu mengisi semua data faktor yang dapat mempengaruhi gangguan tidur. Setelah itu, sistem akan melakukan perhitungan dengan menggunakan algoritma naïve bayes untuk melakukan prediksi. Hasil akan langsung muncul dan menunjukkan apakah pengguna menderita gangguan tidur anemia atau tidak. Data yang digunakan diambil dari website Kaggle yang berisi 374 data klasifikasi dan diambil 12 kelas yang akan digunakan untuk data latih sebagai berikut :

No.	Atribut	Keterangan
1.	JK (Jenis Kelamin/Gender)	1 = Male (Laki-Laki) 2 = Female (Perempuan)
2.	Age (Umur)	Umur.
3.	SD (Sleep Duration)	Rata-rata durasi tidur dalam satuan jam per hari.
4.	QoS (Quality of Sleep)	Level kualitas tidur dengan nilai 1-10.
5.	PAL (Physical Activity Level)	Rata-rata jumlah kilokalori (kcal) yang dibakar per kilogram berat badan per hari.
6.	SL (Stress Level)	Level stress dalam rasio 1-10.
7.	BMI (Body Mass Index)	1 = Normal 2 = Overweight 3 = Obese (obesitas)
8.	SBP (Systolic Blood Pressure)	Tekanan yang dicatat selama kontraksi jantung. Pada pemeriksaan tekanan darah, SBP adalah nilai yang ada di atas.
9.	DBP (Diastolic Blood Pressure)	Tekanan yang dicatat selama relaksasi jantung. Pada pemeriksaan tekanan darah, DBP adalah nilai yang ada di bawah.
10.	HR (Heart Rate)	Rata-rata detak jantung per menit.
11.	DS (Daily Steps)	Langkah Harian.
12.	Sleep Disorder	Gangguan tidur (none, insomnia, sleep apnea).

4. HASIL DAN PEMBAHASAN

Pengembangan sistem prediksi sleep disorder pada penelitian ini berbasis website dengan menggunakan bahasa pemrograman Python dan framework Streamlit.

Halaman Utama

Gambar 4.1 di bawah ini merupakan tampilan Halaman Utama untuk mengisi formulir data dari user yang nantinya akan diprediksi.

Prediksi Sleep Disorder Menggunakan Metode Naive Bayes Gaussian

Aplikasi ini membantu mendeteksi potensi gangguan tidur berdasarkan data kesehatan Anda.

Jenis Kelamin Perempuan	Tekanan Darah Sistolik (mmHg) 130,00	Kategori BMI Normal
Usia 30	Tekanan Darah Diastolik (mmHg) 86,00	Detak Jantung (bpm) 78,00
Durasi Tidur (jam) 6,40	Tingkat Stres (1-10) 7	Langkah Harian 4100
Kualitas Tidur (1-10) 5	Physical Activity Level 35	

Prediksi

Gambar 4. 1 Halaman Utama

Gambar 4.2 di bawah menunjukkan pada halaman utama terdapat fitur untuk menambahkan data latih baru dengan mengunggah dataset yang ingin ditambahkan. Pada tahap ini akan muncul tabel data latih dan data uji yang dipakai dengan keterangan akurasi algoritma Naïve Bayes Gaussian.

Aplikasi ini membantu mendeteksi potensi gangguan tidur berdasarkan data kesehatan Anda.

Unggah file CSV untuk prediksi

Drag and drop file here
Limit 200MB per file • CSV

shl_dataset.csv 16.6KB

Hasil Prediksi dari Data File yang Diunggah

Akurasi pada Data Training: 0.8862876254189692 || Akurasi pada Data Testing: 0.8533333333333334

Data Pelatihan (Training Data)

	Gender	Age	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Systolic Blood Pressure	Diastolic Blood Pressure	Heart Rate	Daily Steps	Sleep
192	1	43	6.5	6	45	7	2	130	85	72	6,000	Inson
75	1	33	6	6	30	8	1	125	80	72	5,000	None
84	1	35	7.5	8	60	5	1	120	80	70	8,000	None
362	2	59	8.2	9	75	3	2	140	95	68	7,000	Sleep

Gambar 4. 2 Halaman Menambah Data Latih Baru

Halaman Hasil

Gambar 4.3 di bawah ini merupakan tampilan Halaman Hasil dari sistem prediksi sleep disorder . Pada halaman hasil ini menampilkan hasil klasifikasi gangguan tidur menggunakan algoritma Naïve Bayes Gaussian. Pada gambar di bawah adalah contoh jika hasil klasifikasi

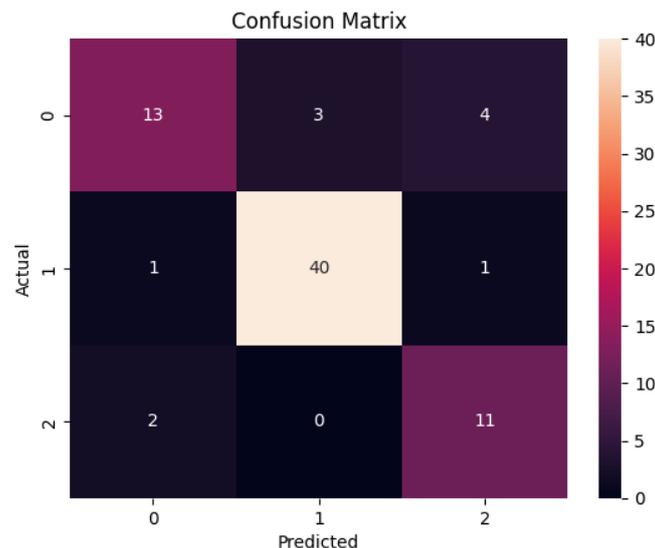
menunjukkan pengguna memiliki gangguan tidur insomnia. Dari halaman ini, pengguna dapat melakukan prediksi ulang dengan klik pada button Prediksi Ulang dan akan diarahkan ke halaman utama.



Gambar 4. 3 Halaman Hasil

Pengujian Sistem

Pengujian Confusion Matrix dilakukan menggunakan modelling python dan diambil 75 data secara acak. Berikut adalah gambar matriks hasil pengujian Confusion Matrix :



Gambar 4. 4 Confusion Matrix

Dari 75 data menunjukkan 64 data dengan prediksi benar dan 11 data dengan prediksi salah. Dari perhitungan Confusion Matrix diatas dengan menggunakan 75 data didapatkan hasil

akurasi 85,3%. Berdasarkan hasil akurasi tersebut, algoritma Naïve Bayes Gaussian tergolong baik untuk prediksi gangguan tidur.

5. KESIMPULAN

Penelitian ini menggunakan dataset dengan jumlah 374 data latih dengan 75 data sebagai data uji. Sesuai dengan hasil analisis pengujian sistem menggunakan Confusion Matrix dengan menguji 75 data yang diambil secara acak didapatkan hasil akurasi sebesar 85,3% . Hal tersebut membuktikan bahwa klasifikasi menggunakan algoritma Naïve Bayes Gaussian tergolong baik untuk memprediksi gangguan tidur. Keterbatasan pada penelitian ini adalah tahap preprocessing yang kurang dilakukan secara mendalam sehingga memungkinkan terjadinya kesalahan pada proses klasifikasi yang mempengaruhi akurasi pada algoritma Naïve Bayes Gaussian. Saran yang dapat penulis berikan untuk penelitian selanjutnya adalah peneliti selanjutnya dapat menambahkan tahap preprocessing yang lebih mendalam dengan harapan mendapat hasil akurasi yang lebih tinggi.

DAFTAR PUSTAKA

- Bollu, P. C., & Kaur, H. (2019). Sleep Medicine: Insomnia and Sleep. *Missouri Medicine*, 116, 68-75.
- Azzahra, S. S. (2019). Obstructive Sleep Apnea (OSA) Sebagai Faktor Resiko Hipertensi. *Jurnal Ilmiah Kesehatan Sandi Husada*, 8(2), 321-324.
- Mubarog, I., Setyanto, A., & Sismoro, H. (2019). Sistem Klasifikasi Pada Penyakit Breast Cancer Dengan Menggunakan Metode Naive Bayes. *Creative Information Technology Journal*, 6(2), 109-118.
- Karim, f., Nurcahyo, G. W., & Sumijan. (2021). Sistem Pakar dalam Mengidentifikasi Gejala Stroke Menggunakan Metode Naive Bayes. *Jurnal Sistem Informasi dan Teknologi*, 3(4), 221-226.
- Sinaga, T. H., Wanto, A., Gunawan, I., Sumarno, S., & Nasution, Z. M. (2021). Implementation of Data Mining Using C4. 5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM. *Journal Of Computing Networks, Architecture and High Performance Computing*, 3(1), 9-20.
- Zai, C. (2022). Implementasi Data Mining Sebagai Pengolahan Data. *Jurnal Portal Data*, 2(3).
- Putri, N., & Wijayanto, A. (2022). Analisis Komparasi Algoritma Klasifikasi Data Mining Dalam Klasifikasi Website Phishing. *Komputika: Jurnal Sistem Komputer*, 11(1), 59-66.

- Alkhairi, P., Zer, Batubara, E., Tambunan, F., & Rosnelly, R. (2021). Pengenalan Pola Kemampuan Pelanggan dalam Membayar Air PDAM Menggunakan Algoritma Naive Bayes. *Jurnal TIMES (Technology Informatic & Computer System)*, 29-38.
- Srirahayu, A., & Pribadie, L. (2023). Review Paper Data Mining Klasifikasi Data Mining. *Jurnal Ilmiah Informatika Global*, 14(1).
- Wahyuningrum, E. (2021). Gangguan Tidur Anak Usia Sekolah. *Jurnal Keperawatan*, 13, 699-708. doi:<https://doi.org/10.32583/keperawatan.v13i3.1387>
- Zulham, M., Saripurna, D., & Siambaton, Z. (2023).). Aplikasi Diagnosa Penyakit Hepatitis dengan Menggunakan Metode Teorema Bayes dan Certainty Factor. *Blend Sains Jurnal Teknik*, 2(1), 1-15.