Jurnal Teknik Mesin, Elektro dan Ilmu Komputer(TEKNIK) Vol.5, No.2 JULI2025

P-ISSN: 2809-9095; E-ISSN: 2809-9125, Hal 281-293 DOI: https://doi.org/10.55606/teknik.v5i2.7609





Diagnosis Gangguan Tidur Berdasarkan Gaya Hidup Menggunakan Algoritma *Naïve Bayes*

Magdalena Herlin Wungubelen

Institut Keguruan dan Teknologi Larantuka

Alfian Nara Weking

Institut Keguruan dan Teknologi Larantuka

Dominikus Boli Watomakin

Institut Keguruan dan Teknologi Larantuka

Alamat: Jl. Kihajar Dewantara Kec. Larantuka - Kab. Flores Timur - Prov. Nusa Tenggara Timur Korespondensi penulis: <u>Ellynwungubelen01@gmail.com</u>

Abstract. Sleep disorders are health problems that often arise due to unhealthy lifestyle patterns and are often overlooked for their impact. This study aims to help detect the risk of sleep disorders using the Naive Bayes algorithm. Data were collected through interviews and examinations, then processed with preprocessing and testing data and achieved a classification accuracy of 88.6% for three categories: Normal, Insomnia, and Sleep Apnea. These results support the application of the Naive Bayes algorithm as a supportive diagnostic method based on lifestyle factors. This finding is also expected to serve as a basis for providing lifestyle improvement recommendations to prevent the risk pf sleep disorders.

Keywords: Sleep Disorder, Lifestyle, Naive Bayes, Classification, Recommendation

Abstrak. Gangguan tidur merupakan masalah kesehatan yang sering muncul akibat pola gaya hidup yang kurang sehat dan sering tidak disadari dampaknya. Penelitian ini bertujuan untuk membantu mendeteksi risiko gangguan tidur menggunakan algoritma Naive Bayes. Data dikumpulkan melalui wawancara dan pemeriksaan, kemudian doproses dengan pra-pemrosesan dan pemyimbangan data menggunakan metode SMOTE. Model diuji dengan pembagian data latih dan data uji, dan menghasilkan akurasi klasifikasi sebesar 88.6% pada tiga kategori, yaitu Normal, Insomnia, dan Sleep Apnea. Hasil ini mendukung penerapan algoritma Naive Bayes sebagai metode diagnosis pendukung berbasis gaya hidup. Temuan ini juga diharapkan dapat menjadi dasar dalam memberikan rekomendasi perbaikan pola hidup guna mencegah risiko gangguan tidur.

Kata kunci: Gangguan Tidur, Gaya Hidup, *Naive Bayes*, Klasifikasi, Rekomendasi Pencegahan

LATAR BELAKANG

Gangguan tidur merupakan salah satu masalah kesehatan yang semakin meningkat di dunia. Menurut Organisasi Kesehatan Dunia (WHO) mencatat bahwa sekitar 10-30% populasi global mengalami gangguan tidur (WHO, 2018). Di Indonesia, hasil Riset Kesehatan Dasar (Riskesdas, 2018) menunjukkan sekitar 12,8% penduduk mengalami gangguan tidur (Kemenkes, 2018). Salah satu faktor utamanya adalah gaya hidup yang tidak sehat, seperti pola makan tidak seimbang, kurangnya aktivitas fisik, kebiasaan merokok, konsumsi kafein yang berlebihan, serta penggunaan perangkat elektronik sebelum tidur merupakan faktor yang dapat memperburuk kualitas tidur seseorang Kurangnya kualitas tidur yang baik dapat memicu berbagai masalah kesehatan, termasuk peningkatan risiko penyakit kardiovaskular, obesitas, diabetes, dan gangguan mental seperti depresikecemasan serta (permana et al., 2023).

Deteksi dini gangguan tidur sangat penting untuk mencegah dampak negatifnya. Salah satu pendekatan yang dapat digunakan adalah penerapan algoritma klasifikasi seperti *Naive Bayes* (Nisa, 2024). Penelitian ini mengevaluasi efektivitas algoritma *Naive Bayes* dalam mendiagnosis gangguan tidur berdasarkan gaya hidup, mengidentifikasi pola gaya hidup yang berkontribusi, serta memberikan rekomendasi perbaikan.

KAJIAN TEORITIS

a. Diagnosis

Diagnosis adalah proses identifikasi dan penentuan suatu kondisi atau penyakit berdasarkan gejala dan tanda-tanda yang ada. Dalam konteks judul "Diagnosis Gangguan Tidur Berdasarkan Gaya Hidup Menggunakan Algoritma Naïve Bayes," diagnosis mengacu pada proses penentuan jenis gangguan tidur dengan menganalisis data gaya hidup menggunakan metode algoritma klasifikasi(Adolph 2016).

b. Gangguan Tidur

Gangguan tidur adalah kondisi yang mengubah pola tidur seseorang dan berdampak buruk pada kesehatan dan kualitas hidup (Radityo, 2012). Gejala umum meliputi sulit tidur, sering terbangun, mudah lelah, dan pola tidur tidak teratur. Jika tidak diatasi, gangguan tidur meningkatkan risiko penyakit serius seperti hipertensi dan jantung. Beberapa jenis gangguan tidur meliputi: Insomnia (kesulitan tidur, bangun terlalu pagi; Fernando & Hidayat, 2020), Sleep Apnea (henti napas sejenak saat tidur; Basuki Rahman, Handoyo, & Rohadi, 2012), Narcolepsy (kantuk berlebihan dan serangan tidur mendadak; Zaini & Manesh, 2020), serta Restless Legs Syndrome/RLS (dorongan menggerakkan kaki karena sensasi tidak nyaman; Surya & Desnita, 2019).

c. Gaya Hidup

Gaya hidup adalah pola perilaku sehari-hari (KBBI). Gaya hidup sehat mencegah penyakit kronis dan meningkatkan kualitas hidup (WHO, 2018). Aspeknya meliputi aktivitas fisik, kualitas tidur, durasi tidur, tingkat stres, tekanan darah, detak jantung, langkah harian meningkatkan resiko penyakit jantung (AHA, 2017). Gaya hidup sehat bermanfaat bagi kesehatan (Fitriani, 2021).

d. Naive Bayes

Metode Naive Bayes adalah teknik klasifikasi berbasis probabilistik yang menggunakan Teorema Bayes dengan asumsi independensi antar fitur.(Devita, Herwanto, and Wibawa 2018)

e. Penelitian Terdahulu

- (Iwan Ady Prabowo, Dwi Remawati dan Aji Pratama Wisnu Wardana, 2020): Klasifikasi Tingkat Gangguan Tidur Menggunakan Algoritma Naive Bayes, dengan tingkat akurasinya 80%.
- 2. (Juan Perez, 2020): Menggunakan Metode Dempster Shafer untuk mendiagnosa gangguan tidur, dengan tingkat akurasinya 90%.

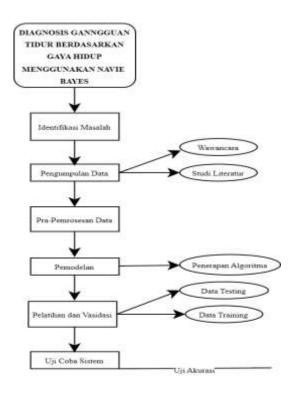
- 3. (Ahmad Revaldo, Yupianti dan Ila Yati Beti, 2013): Menggunakan Metode Forward Chaining Berbasis Web untuk Sistem Pakar Diagnosa Penyakit Gangguan Tidur, dengan tingkat akurasinya 100%.
- 4. (Diana Sari, 2024): Menggunakan Support Vector Machine dan Neura Network untuk memprediksi Gangguan Tidur pada Sleep Health dan Lifestyle, dengan tingkat akurasinya 90.1% untuk Support Vector Machine danalgoritma Neural Networkpersentasenya 91.2%.
- 5. (Mohamad Reyhand Fatturrahman dan AliyahKurniasih, 2023): KlasifikasiGangguanTidurBerdasarkanKualitasTidurdanGayaHidup menggunakan MetodeNearMiss,SMOTE,danNaïveBayes, dengan tingkat akurasinya 81%

Berdasarkan kajian teori dan penelitian terdahulunya yang telah dilakukan oleh beberapa peneliti dengan metode klasifikasi yang berbeda, penelitian ini dilaksanankan untuk mendiagnosis gangguan tidur dengan mempertimbangkan faktor gaya hidup, menggunakan algoritma *Naive Bayes* sebagai metode klasifikasi utama yang diharapkan dapat meningkatkan akurasi hasil klasifikasi yang lebih baik, sehingga hasil diagnosis gangguan tidur yang dihasilkan lebih optimal.

METODE PENELITIAN

Penelitian ini merupakan penelitian terapan yang menggunakan algoritma *Naive Bayes* untuk mendiagnosis gangguan tidur dengan mempertimbangkan pola gaya hidup. Penelitian ini terdiri dari beberapa tahapan yang dirangkum dalam proses "Penerapan Algoritma Naive Bayes untuk Diagnosis Gangguan Tidur Berdasarkan Gaya Hidup".

Adapun diagram alur penelitiannya sebagai berikut:



Gambar 1. Metode Penelitian

Dengan menggunakan pendekatan kuantitatif eksperimental, penelitian ini terdiri dari serangkaian tahapan, yaitu pengumpulan data responden mengenai gaya hidup dan kualitas tidur, tahapan praproses (cleaning, normalisasi, transformasi), ekstraksi fitur utama, pembagian data ke data latih dan uji, penerapan algoritma *Naive Bayes*, validasi model melalui *cross-validation*, serta evaluasi performa model klasifikasi. Penelitian ini bertujuan untuk mrngeksplorasi sejauh mana algoritma *Naive bayes* mampu menghasilkan diagnosis gangguan tidur yang akurat berdasarkan pola gaya hidup individu.

HASIL DAN PEMBAHASAN

1. Dataset

Penelitian ini menggunakan dua sumber data. Data sekunder diambil dari platform Kaggle, yaitu "Sleep Health and Lifestyle Dataset" (https://www.kaggle.com/datasets/ajay1735/sleep-health-and-lifestyle-dataset) dengan total 373 data responden yang memuat atribut usia, jenis kelamin, durasi tidur, kualitas tidur, tingkat aktivitas fisik, tingkat stres, tekanan darah, detak jantung, langkah harian, dan gangguan tidur. Selain itu, data primer sebanyak 10 data

dikumpulkan langsung dari UPT Puskesmas Nagi pada periode 14 Juni-14 Juli 2025 melalui wawancara dan pemeriksaan fisik oleh tim medis. Data primer digunakan sebagai data uji untuk memvalidasi model prediksi gangguan tidur pada kondisi nyata.

Person ID	Jenis Kelamin	Usia	Durasi Tidu r	Kualitas Tidur	Tingkat Aktivitas Tidur	Tingka t Stres	Kategori IMT	Tekanan Darah	Detak Jantung	Langkah harian	Gangguan Tidur
1	Pria	27	61	6	42	6	Kelebihan BB	12683	77	4200	Normal
2	Pria	28	62	6	60	8	Normal	12580	75	10000	Normal
3	Pria	28	62	6	60	8	Normal	12580	75	10000	normal
4	Pria	28	59	4	30	8	Kegemukan	14090	85	3000	Slep apmea
	Pria	28	59	4	30	8	Kegemukan	14090	85	3000	Slep Apmea
373	Pria	28	59	4	30	8	Kegemukan	14090	85	3000	Insomia

Tabel 1. Data Sekuder

Person ID	Jenis Kelamin	Usia	Durasi Tidur	Kualitas Tidur	Tingkat Aktivitas Tidur	Tingkat Stres	Tekanan Darah	Detak Jantung	Langkah harian	Gangguan Tidur
1	Permpuan	25	5.5	5	30	8	120/70	78	3000	Insomnia
2	Pria	41	6.8	7	50	6	140/88	88	6000	Sleep Apnea
3	Perempuan	32	7.2	9	80	3	118/72	72	8500	Normal
4	Pria	28	6.1	6	40	7	130/82	82	4000	Insomnia
5	Perempuan	55	7	7	45	6	142/9	90	5000	Sleep Apnea
6	Pria	39	7.4	8	75	2	125/70	70	9000	Normal
7	Perempuan	45	6.3	5	25	8	135/85	85	3500	Insomnia
8	Pria	36	6.9	7	55	5	138/86	86	6200	Sleep Apnea
9	Perempuan	30	7.3	9	85	3	122/75	75	8700	Normal
10	Pria	48	5.9	5	20	9	130/82	82	2800	Insomnia

Tabel 2. Data Primer

2. DataPreprocessing

Tahap pra-pemrosesan data dilakukan untuk memastikan bahwa data yang digunakan dalam proses klasifikasi memiliki kualitas yang baik dan siap untuk dianalisis oleh algoritma Naïve Bayes, Pada tahap ini, dilakukan beberapa langkah penting, antara lain :

a. Pembersihan Data (Data Cleaning)

Pada tahap ini, pembersihan data adalah proses menghilangkan ataumemperbaiki data yang dianggap tidak valid, tidak lengkap, duplikat, atau tidak relavan agar hasil analisis menjadi akurat dan dapat dipercaya.

1) Menghapus Duplikat Data

Penghapusan duplikat merupakan tahap awal data cleaning untuk mencegah pembobotan berlebihan pada data yang sama. Dari 373 data, sebanyak 224 data teridentifikasi sebagai duplikat, sehingga data akhir yang digunakan sebanyak 133 data.

Person ID	Jenis Kelamin	Usia	Durasi Tidu r	Kualitas Tidur	Tingkat Aktivitas Tidur	Tingka t Stres	Kategori IMT	Tekanan Darah	Detak Jantung	Langkah harian	Gangguan Tidur
1	Pria	27	61	6	42	6	Kelebihan BB	12683	77	4200	Normal
2	Pria	28	62	6	60	8	Normal	12580	75	10000	Normal
3	Pria	28	59	4	30	8	Kegemukan	14090	85	3000	Slep Apmea
4	Pria	28	59	4	30	8	Kegemukan	14090	85	3000	Insomnia
	Pria	29	63	6	40	7	Kegemukan	14090	82	3500	Insomnia
373	Pria	29	78	7	75	6	normal	14080	70	8000	Normal

Tabel 3. Hasil Duplikat Data

2) Menghapus Atribut

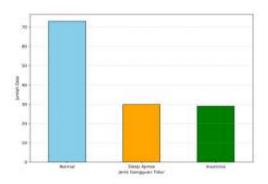
Setelah menghapus duplikat data, langkah selanjutnya adalah menghapus atribut yang tidak penting adalah fitur dalam dataset yang tidak memilki nilai informatif untuk analisis atau model yang akan dibangun. Hal ini untuk memilih dan mempertahankan hanya atribut yang relavan. Atribut yang akan dihapus adalah "Kategori IMT".

Person ID	Jenis Kelamin	Usia	Durasi Tidu r	Kualitas Tidur	Tingkat Aktivitas Tidur	Tingka t Stres	Kategori IMT	Tekanan Darah	Detak Jantung	Langkah harian	Gangguan Tidur
1	Pria	27	61	6	42	6	Kelebihan BB	12683	77	4200	Normal
2	Pria	28	62	6	60	8	Normal	12580	75	10000	Normal
3	Pria	28	59	4	30	8	Normal	14090	85	3000	Slep Apmea
4	Pria	28	59	4	30	8	Kegemukan	14090	85	3000	Insomnia
	Pria	29	63	6	40	7	Kegemukan	14090	82	3500	Insomnia
373	Pria	29	78	7	75	6	Kegemukan	12080	70	8000	Normal

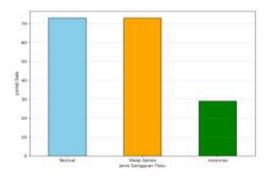
Tabel 4. Penghapusan Atribut

3) Menyimbang Data

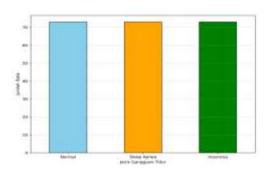
Pada tahap ini, data diseimbangkan menggunakan SMOTE (Synthetic Minority Over-Sampling Technique). Sebelum penyimbangan terdapat 73 data Normal, 30 Sleep Apnea, dan 29 Insomnia. SMOTE menambah sampel sintetis pada kelas minoritas hingga setara dengan kelas mayoritas, sehingga distribusi data menjadi seimbang dan model yang dibangun tidak bias terhadap kelas tertentu, sehingga data yang awalnya 133 akan bertambah menjadi 219 data.



Gambar 2. Diagram batang, data sebelum dilakukan SMOTE



Gambar 3. hasil SMOTE data Sleep Apnea



Gambar 4. hasil SMOTE data Insomnia

Kelas	Sebelum SMOTE	Sesudah SMOTE		
Normal	73	73		
Insomnia	29	73		
Sleep Anea	30	73		
Total	133	219		
Akurasi	72,3%	88,6%		
Keterangan	Data tidak seimbang, prediksi bias ke kelas mayoritas	Data seimbang, prediksi merata pada tiap kelas		

Tabel 5.Distribusi data dan Perbandingan akurasi sebelum dan sesudah SMOTE

Tabel diatas menunjukkan bahwa sebelum SMOTE, data tidak seimbang dan akurasi model hanya 72,3% karena bias ke kelas mayoritas. Setelah SMOTE, jumlah tiap kelas setara 73 sampel, total data naik menjadi 219, dan akurasi meningkat menjadi 88.6% dengan prediksi yang lebih merata.

3. Data Spliting

Setelah preprocesing data hasil SMOTE sebanyak 219 sampel (73 Normal, 73 Insomnia, dan 73 Sleep Apnea) dibagi menggunakan *train_test_split* (scikit-learn) dengan rasio 80% data latih dan 20% data uji. Hasilnya diperoleh 175 data latih (58 per kelas) dan 44 data uji (15 per kelas)

Label	Absolute Count	Fraction		
Insomnia	59	0.337		
Normal	58	0.331		
Sleep Apnea	58	0.331		

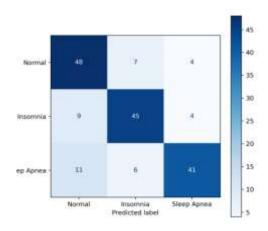
Tabel 6. Jumlah data latih tiap label

Label	Absolute Count	Fraction		
Insomnia	14	0.318		
Normal	15	0.341		
Sleep Apnea	15	0.341		

Tabel 7. Jumlah data uji tiap label

4. Model Building

Tahap selanjutnya adalah pembangunan model menggunakan 80% data latih dengan algoritma Naive Bayes. Berdasarkan confusion matrix, model berhasil mengklasifikasikan 48 data Normal, 45 Insomnia, dan 41 Sleep Apnea dengan benar, meski masih terdapat beberapa kesalahan prediksi, seperti 11 data Sleep Apnea yang terklasifikasi sebagai Normal.



Gambar 5. Confusion Matrix Data Latih

5. Model Validation

Model Naïve Bayes pada penelitian ini divalidasi menggunakan 10 data primer yang dikumpulkan langsung dari UPT Puskesmas Nagi. Hasil validasi menunjukkan bahwa model mampu memprediksi gangguan tidur pada data nyata dengan tingkat akurasi 100% dan distribusi prediksi yang sesuai pada setiap kelas gangguan tidur. Capaian ini membuktikan bahwa model yang dibangun dapat diterapkan secara langsung untuk mendukung diagnosis risiko gangguan tidur berdasarkan indikator gaya hidup dan kondisi kesehatan responden di lapangan.

Person ID	Jenis Kelami n	Usia	Durasi Tidur	Kualitas Tidur	Tingkat Aktivitas Tidur	Tingkat Stres	Tekanan Darah	Detak Jantung	Langkah harian	Gangg uan Tidur	Prediksi
1	Permpu an	25	5.5	5	30	8	120/70	78	3000	Insomn ia	Insomni a
2	Pria	41	6.8	7	50	6	140/88	88	6000	Sleep Apnea	Sleep Apnea
3	Peremp uan	32	7.2	9	80	3	118/72	72	8500	Normal	Normal
4	Pria	28	6.1	6	40	7	130/82	82	4000	Insomn ia	Insomni a
5	Peremp uan	55	7	7	45	6	142/9	90	5000	Sleep Apnea	Sleep Apnea
6	Pria	39	7.4	8	75	2	125/70	70	9000	Normal	Normal

7	Peremp uan	45	6.3	5	25	8	135/85	85	3500	Insomn ia	Insomni a
8	Pria	36	6.9	7	55	5	138/86	86	6200	Sleep Apnea	Sleep Apnea
9	Peremp uan	30	7.3	9	85	3	122/75	75	8700	Normal	Normal
10	Pria	48	5.9	5	20	9	130/82	82	2800	Insomn ia	Insomni a

Tabel 8. Perbandingan Label Aktual dan Hasil Prediksi Model pada Data Lapangan

Kelas	Precision	Recall	F1-Score
Normal	100%	100%	100%
Insomnia	100%	100%	100%
Sleep Apnea	100%	100%	100%
Rata-rata	100%	100%	100%

Tabel 9. Hasil Evaluasi Pada Validasi Data Lapangan.

KESIMPULAN DAN SARAN

Penelitian ini membuktikan bahwa algoritma Naive Bayes mampu melakukan klasifikasi gangguan tidur dengan memanfaatkan pola gaya hidup dan menghasilkan tingkat akurasi yang memuaskan. Dengan menerapkan teknik balancing data melalui SMOTE serta membagi data uji sebesar 20% model berhasil meraih akurasi sebesar 88.6% dan prediksi yang merata pada kelas Normal, Insomnia, dan Sleep Apnea. Temuan ini mengindikasikan bahwa Naive Bayes efektif dalam mengenali pola data dan berpotensi digunakan untuk mendukung proses diagnosis gangguan tidur berbasis atribut gaya hidup.

DAFTAR REFERENSI

Adolph, Ralph. 2016. "済無No Title No Title No Title." 1–23.

Basuki Rahman, Usep, Handoyo, and Pujo Rohadi. 2012. "Hubungan Obesitas Dengan Risiko Obstructive Sleep Apnea (Osa) Pada Remaja." *Jurnal Ilmiah Kesehatan Keperawatan* 8(1):44–56.

- Devita, Riri Nada, Heru Wahyu Herwanto, and Aji Prasetya Wibawa. 2018. "Perbandingan Kinerja Metode Naive Bayes Dan K-Nearest Neighbor Untuk Klasifikasi Artikel Berbahasa Indonesia." *Jurnal Teknologi Informasi Dan Ilmu Komputer* 5(4):427–34.
- Fernando, Renaldo and Ridha Hidayat. 2020. "Hubungan Lama Penggunaan Media Sosial Dengan Kejadian Insomnia Pada Mahasiswa Fakultas Ilmu Kesehatan Universitas Pahlawan Tuanku Tambusai Tahun 2020." *Jurnal Ners* 4(2):84.
- FITRIANI, D. 2021. "Pengaruh Gaya Hidup Dan Literasi Keuangan Terhadap Perilaku Konsumtif Belanja Online Pada Generasi Milenial Di Kabupaten" *Jurnal Riset Sains Manajemen* 2(3):103–10.
- Khasanah, Nurul, Daniati Uki, Eka Saputri, Faruq Aziz, and Taopik Hidayat. 2025. "Studi Perbandingan Algoritma Random Forest Dan K-Nearest Neighbors (KNN) Dalam Klasifikasi Gangguan Tidur." 5(1):17–25.
- Nisa, Indah Chairun. 2024. "Penerapan Algoritma K-Means Dalam Pengkategorian Insomnia." *METHOMIKA Jurnal Manajemen Informatika Dan Komputerisasi Akuntansi* 8(1):117–22.
- Permana, Akbar Satria, Rebecca Sihite, Karenina Safitri, and Muhammad Raihan Aly. 2023. "Hubungan Pola Tidur Terhadap Potensi Penyakit Diabetes Di Usia Muda Pada Mahasiswa Universitas Negeri Semarang." *Jurnal Analis* 2(2):115–23.
- Radityo, Wyn Eko. 2012. "Depresi Dan Gangguan Tidur." *E-Jurnal Medika Udayana* 1 (1):1–16.
- Surya, Defrima Oka and Ria Desnita. 2019. "Efektifitas Akupresur Terhadap Derajat Restless Leg Syndrome Pada Pasien Yang Menjalani Hemodialisis." *JURNAL KESEHATAN PERINTIS (Perintis's Health Journal)* 6(2):104–9.
- Zaini, Mimi Fitriana and Niloufar Heshmati Manesh. 2020. "The Impact of Dance on the Development of Coping Mechanisms Fornarcolepsy: A Narrative Analysis." *Jurnal Ilmiah Peuradeun* 8(1):17–36.