Content lists available at LiteraPublishing
Journal of Public Health and Community Systems
NEXURA
journal homepage : https://journal.literapublishing.id/index.php/nexura

Article

Sleep Disorder Classification Using the Naïve Bayes Algorithm

Olivia Natania Anjani [1,*], Khusnul Fatimah Azzahra [2], Anisa Fitria Rahma [3], Aida Rahma Saqina [4], Muhammad Yunus [5]

[1] Health Information Management, Politeknik Negeri Jember, Jember, Indonesia; nataniaolivia64@gmail.com
[2] Health Information Management, Politeknik Negeri Jember, Jember, Indonesia; khusnulfatimahazzahra288@gmail.com
[3] Health Information Management, Politeknik Negeri Jember, Jember, Indonesia; anisafitriarahma04@gmail.com
[4] Health Information Management, Politeknik Negeri Jember, Jember, Indonesia; aidarahma471@gmail.com
[5] Health Information Management, Politeknik Negeri Jember, Jember, Indonesia; m.yunus@polije.ac.id

* Correspondence: nataniaolivia64@gmail.com

Citation: O. N. Anjani, K. F. Azzahra, A. F. Rahma, A. R. Saqina, and M. Yunus, "Sleep Disorder Classification Using the Naïve Bayes Algorithm", *NEXURA*, vol. 1, no. 1, pp. 1–11, Jun. 2026.

Received: 26-05-2026
Accepted: 20-06-2026
Published: 30-06-2026

Abstract: Sleep disorders are a medical condition with a continuously increasing prevalence globally, including in Indonesia, and have a serious impact on an individual's quality of life, cognitive function, and cardiovascular health. This study aims to classify types of sleep disorders using the Naïve Bayes algorithm as a probabilistic approach in health data mining. The dataset used is the Sleep Health and Lifestyle Dataset, consisting of 400 data records with 13 attributes, covering demographic factors, lifestyle, and physiological parameters. The research stages included data preprocessing, blood pressure feature engineering, and splitting the data into training and test sets with an 80:20 ratio. Model evaluation was performed using a confusion matrix and accuracy, precision, recall, and F1-score metrics. The results showed that the Naïve Bayes model achieved an accuracy of 71.60% on the 81-sample test set. The best performance was observed in the "None" class with an F1-score of 0.83, while the "Insomnia" and "Sleep Apnea" classes could not be successfully identified by the model due to the dominance of the majority class (class imbalance). An example prediction for a 47-year-old female subject with a sleep quality of 7/10, a stress level of 4/10, and a Normal BMI resulted in a classification of "No Sleep Disorder" with a probability of 72.08%. These findings indicate that addressing class imbalance is necessary to improve the classification performance of the Naïve Bayes model on the sleep disorder dataset.

Keywords: Classification; Data Mining; Naïve Bayes, Prediction; Sleep Disorder

1. Introduction

Sleep is a fundamental biological need that plays a role in physical recovery, emotional regulation, memory consolidation, and strengthening the body's immune system. When these vital functions are disrupted, the impact is not only felt physiologically but also affects an individual's overall quality of life. Conditions that hinder sleep function are known as sleep disorders, defined as pathological conditions that disrupt normal sleep patterns, quality, or duration, thereby negatively impacting an individual's functioning during waking hours [1].

One of the most common yet frequently undiagnosed forms of sleep disorder is Obstructive Sleep Apnea (OSA). OSA is characterized by recurrent episodes of apnea and hypopnea during sleep, with a wide spectrum of severity ranging from mild to severe. Population-based studies indicate that even mild OSA has been associated with various health complications, including hypertension, cardiovascular disease, stroke, excessive daytime sleepiness, motor vehicle accidents, and reduced quality of life [2]. If left untreated, this condition carries a risk of triggering cardiac arrhythmias and various other complications that contribute to increased morbidity and mortality.

In addition to OSA, insomnia is also a sleep disorder with a very high prevalence globally. Based on a recent systematic review analyzing data from 18 studies involving a total of 262,582 participants, the prevalence of insomnia is estimated to reach 16.2% in the global adult population, equivalent to more than 852 million individuals [3]. In Indonesia itself, a study using data from the 2014–2015 Indonesia Family Life Survey (IFLS-5) found that the prevalence of clinical insomnia among residents aged 15 years and older reached 11.0%. The study also identified several factors associated with the occurrence of insomnia, including low socioeconomic status, low educational attainment, chronic health conditions, and certain health behaviors [4]. This high prevalence underscores that sleep disorders are not merely individual issues but have become a public health problem requiring systemic attention.

The impact of sleep disorders on health is multidimensional. Prolonged insomnia is associated with an increased risk of cardiovascular disorders, type 2 diabetes, depression, anxiety, and a decline in overall quality of life [5]. In addition to these clinical impacts, sleep disorders also impose a significant economic burden through reduced work productivity and increased utilization of healthcare services [3]. Given the magnitude of these impacts, early detection of sleep disorders is crucial. However, some types of sleep disorders share similar symptoms, making manual identification time-consuming and prone to errors [6]. This situation has driven the development of information technology-based systems to help healthcare professionals analyze and classify sleep disorders more efficiently.

In this context, the application of classification techniques in data mining has yielded many promising results in the field of healthcare. Various classification algorithms, such as Naïve Bayes, have been successfully utilized to aid in the diagnosis and prediction of diseases, including hepatitis C and breast cancer [7], [8].Among the various available algorithms, Naïve Bayes is a probabilistic method known for its simple computational process yet still delivering competitive classification performance. This algorithm operates based on Bayes' theorem under the assumption of conditional independence among features. Although this assumption is rarely fully met in real-world data, Naïve Bayes has proven to perform well on various problems, including in the healthcare field [9], [10].

Several previous studies have explored the application of classification to sleep disorders. Prabowo et al. (2020) successfully classified several types of sleep disorders, such as insomnia, hypersomnia, narcolepsy, sleep terrors, sleep schedule disorders, and nightmares, using the Naïve Bayes algorithm with an accuracy rate of 80% [6]. Research by Sari et al. (2025), which utilized sleep disorder cases based on data from the National Institute of Neurological Disorders and Stroke using various algorithms, showed that Naïve Bayes contributed to performance comparisons, while ensemble algorithms such as Random Forest and Decision Tree generally achieved higher accuracy but with significantly greater model complexity [11]. Meanwhile, Faradillah et al. (2024) developed a machine learning-based sleep disorder prediction model by considering risk factors such as age, sleep duration, BMI category, and occupation, where the Gradient Boosting model delivered the best performance with an accuracy of 99% [12].

Nevertheless, research specifically focused on the application of Naïve Bayes with a comprehensive lifestyle dataset, while evaluating it from a health information management perspective, remains very limited. Based on this research gap, this study aims to: (1) analyze the characteristics of the Sleep Health and Lifestyle Dataset as a basis for classification; (2) apply the Naïve Bayes algorithm in the process of classifying types of sleep disorders; and (3) evaluate model performance using a confusion matrix and standard evaluation metrics. The results of this study are expected to contribute to the development of data mining-based clinical decision support systems in the context of health information management.

2. Materials and Methods

2.1 Type of Research

The Materials and Methods should be described with sufficient details to allow others to replicate and build on the published results. Please note that the publication of your manuscript implicates that you must make all materials, data, computer code, and protocols associated with the publication available to readers. Please disclose at the submission stage any restrictions on the availability of materials or information. New methods and protocols should be described in detail while well-established methods can be briefly described and appropriately cited.

Research manuscripts reporting large datasets that are deposited in a publicly available database should specify where the data have been deposited and provide the relevant accession numbers. If the accession numbers have not yet been obtained at the time of submission, please state that they will be provided during review. They must be provided prior to publication. Interventionary studies involving animals or humans, and other studies that require ethical approval, must list the authority that provided approval and the corresponding ethical approval code.

2.2 Dataset and Data Sources

The dataset used is the Sleep Health and Lifestyle Dataset (https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset), which is publicly available via the Kaggle platform. This dataset consists of 400 data records representing the health and lifestyle profiles of individuals with various demographic and clinical backgrounds. The same dataset has been used in studies related to the prediction of sleep disorders using machine learning methods [11], [12], making it a valid and reliable reference for similar research.

2.3 Research Variables/Attributes

The dataset consists of 12 predictor attributes and 1 target attribute (Sleep Disorder). Table 1 presents a complete description of each variable used in this study.

Table 1. Research Dataset Variables and Attributes

Attribute	Data Type	Role	Description
Gender	Categorical	Predictor	Sex (Male/Female)
Age	Numeric	Predictor	Age in years
Occupation	Categorical	Predictor	Job type (4 categories)
Sleep Duration	Numeric	Predictor	Average sleep duration (hours/night)
Sleep Quality	Numeric	Predictor	Sleep quality (scale 1–10)
Physical Activity Level	Numeric	Predictor	Physical activity (minutes/day)
Stress Level	Numeric	Predictor	Stress level (scale 1–10)
BMI Category	Categorical	Predictor	BMI Category (4 classes)
Blood Pressure (Systolic)	Numeric	Predictor	Systolic blood pressure (mmHg)
Blood Pressure (Diastolic)	Numeric	Predictor	Diastolic blood pressure (mmHg)
Heart Rate	Numeric	Predictor	Heart rate (bpm)
Daily Steps	Numeric	Predictor	Number of steps per day
Sleep Disorder	Categorical	Target	None / Insomnia / Sleep Apnea

2.4 Data Preprocessing

The data preprocessing stage is a critical step prior to applying classification algorithms to ensure data quality, consistency, and readiness [13]. By performing data preprocessing correctly, we can ensure that the data used in analysis or modeling is cleaner, more relevant, and more reliable for deriving valuable insights and information from the collected data. The purpose of data preprocessing is to clean, prepare, and transform the data into a suitable format so that it is easier to process, understand, and interpret by the analytical tools or models to be used [14]. The preprocessing conducted in this study includes the following four stages:

1. Checking and handling missing values: All attributes were checked to identify and handle incomplete values so they would not interfere with the model training process.
2. Separation of blood pressure features: The Blood Pressure attribute, which was originally in a combined systolic/diastolic format, was split into two independent numerical columns.
3. Dataset splitting (train-test split): The dataset was divided into a training set of 80% (319 data points) and a test set of 20% (81 data points) using the holdout method.

2.5 Naïve Bayes Algorithm

Naïve Bayes is a probabilistic classification method based on Bayes' Theorem, assuming that each attribute is independent of the others [6]. This algorithm falls under supervised learning methods because its training process uses training data that has been pre-labeled with class labels. This assumption of independence among attributes is why the method is known as the naïve or simple method [15]. Despite its strong assumptions, Naïve Bayes remains widely used because its training process is simple, efficient, and capable of delivering good performance even with a relatively small amount of training data [6]. In the Naïve Bayes method, each variable is considered independent of the others, so the calculation process only requires estimating the parameters of each variable for each class. With this assumption, the computational complexity is lower compared to other classification methods because there is no need to calculate the relationships among all variables [6].

Bayes' Theorem is used to determine the probability of a class based on known data or attributes. The Naïve Bayes algorithm combines the prior probability with the conditional probability to produce the posterior probability, which serves as the basis for classification decision-making. The general form of Bayes' Theorem can be written as follows [6]:

$$P(C|X) = P(X|C) \times P(C) / P(X)$$

where:

- X = data or attributes whose class is unknown.
- H = the hypothesis that data X belongs to a specific class.
- $P(H|X)$ = posterior probability, i.e., the probability of hypothesis H after data X is known.
- $P(H)$ = prior probability, i.e., the initial probability of a class.
- $P(X|H)$ = likelihood probability, i.e., the probability of data X occurring in class H.
- $P(X)$ = evidence probability, which is the overall probability of data X.

Based on the posterior probability obtained, the algorithm will determine the class with the highest probability value as the final classification result.

2.6 Model Evaluation

Model performance evaluation is conducted using a confusion matrix that summarizes prediction results in a multi-class classification matrix [16]. The evaluated metrics include:

(1) Accuracy: the proportion of correct predictions out of the total test data;
(2) Precision: the proportion of true positives among positive predictions per class;
(3) Recall: the proportion of positive instances successfully identified by the model; and
(4) F1-score: the harmonic mean of precision and recall, which is useful under conditions of imbalanced class distribution [16].

3. Results and Discussion

3.1 Dataset Characteristics

The results of the descriptive analysis of the Sleep Health and Lifestyle Dataset show the distribution of respondent characteristics as presented in Table 2. The dataset consists of 400 data records with a relatively balanced gender distribution (49.75% male and 50.25% female). The occupational groups consist of Students (27.5%), Office Workers (24.75%), Retirees (23.75%), and Manual Laborers (24%). The BMI category distribution includes Overweight (27.25%), Underweight (25.5%), Obese (24.5%), and Normal (22.75%).

Table 2. Distribution of Sleep Disorder Classes in the Dataset

Sleep Disorder Class	Count	Percentage
None (No Disorder)	290	72.50%
Insomnia	79	19.75
Sleep Apnea	31	7.75%
Total	**400**	**100%**

3.2 Classification System Interface

Figure 1 shows the interface of the sleep disorder classification system, designed to display the prediction process using the Naïve Bayes method. This system allows users to enter patient data, including demographic data such as gender, age, and occupation; sleep data such as sleep duration, sleep quality, and stress levels; health data such as BMI category, systolic and diastolic blood pressure, and heart rate; and physical activity data consisting of physical activity and daily step count. After all data is entered, the system performs the classification process using the Naïve Bayes algorithm and displays the predicted sleep disorder results along with the probability values for each class.

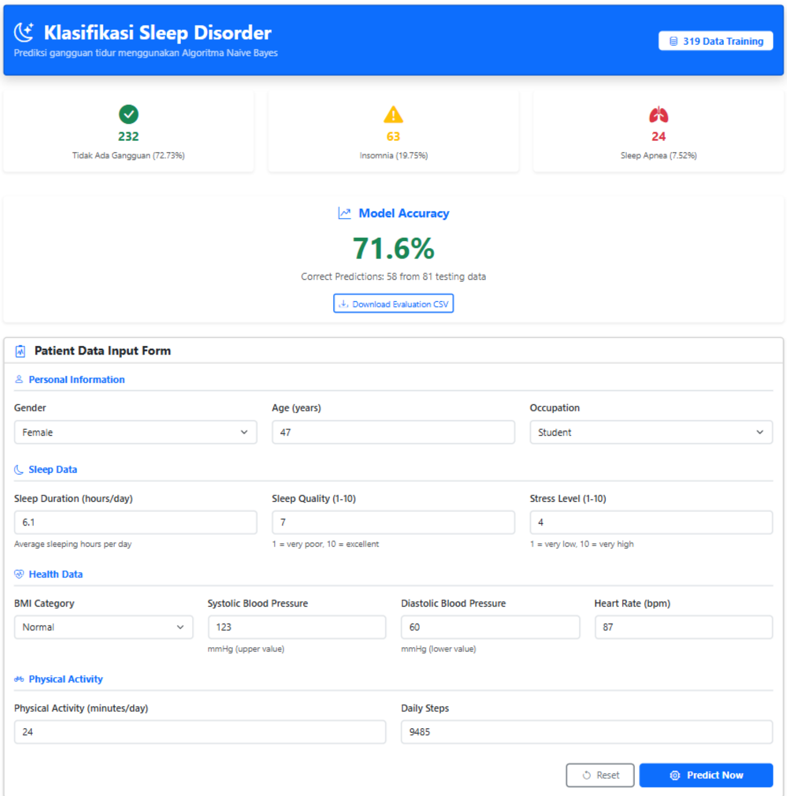

Figure 1. Patient Data Input Form Interface in the Sleep Disorder Classification System

3.3 *Example of Classification Results*

Figure 2 shows an example of classification results generated by the system for a 47-year-old female patient who is a student. The entered data includes a sleep duration of 6.1 hours per day, sleep quality of 7 on a scale of 10, stress level of 4 on a scale of 10, Normal BMI category, blood pressure of 123/60 mmHg, heart rate of 87 bpm, physical activity of 24 minutes per day, and a daily step count of 9,485 steps.

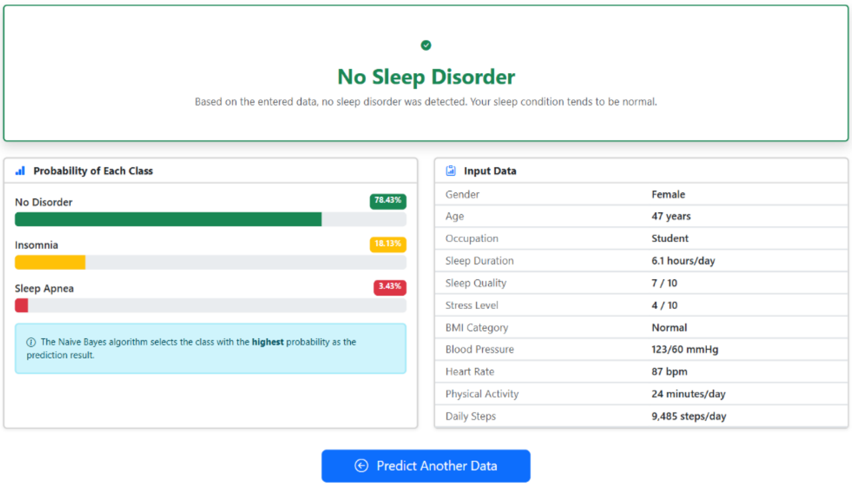

Figure 2. Display of Sleep Disorder Classification Results by the Naïve Bayes-Based System

Based on the entered data, the system classified the patient into the **No Sleep Disorder** category with the highest probability value of 80.00%. Meanwhile, the probability for the **Insomnia** category was 15.00% and for **Sleep Apnea** was 5.00%.

These results align with the patient's condition, which includes fairly good sleep quality (7/10), low stress levels (4/10), and a Normal BMI category. However, the patient's sleep duration of 6.1 hours per day remains slightly below the recommended ideal sleep duration for adults, which is approximately 7–9 hours per day.

3.4 Model Evaluation Using a Confusion Matrix

Model testing was conducted on 80 test data points (20% of the total dataset). The test results are presented in the confusion matrix in Table 3 and summarized in the evaluation metrics in Table 4.

Table 3. Confusion Matrix of Naïve Bayes Model Classification Results

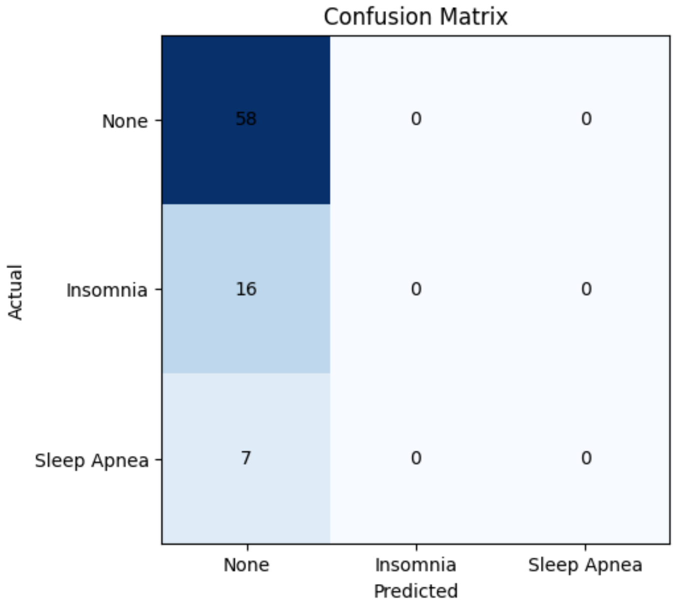

Table 4. Naïve Bayes Model Evaluation Metrics by Class

Class	Precision	Recall	F1-Score	Support
None (No Interference)	0.72	1.00	0.83	58
Insomnia	0.00	0.00	0.00	16
Sleep Apnea	0.00	0.00	0.00	7
Overall Accuracy	**—**	**—**	**71.0**	**81**

Based on Table 3 and Table 4, the Naïve Bayes model achieved an accuracy of 71.60% on 81 test data points. The best performance was demonstrated in the None class with a precision of 0.72, a recall of 1.00, and an F1-score of 0.83. However, the model's performance in the Insomnia and Sleep Apnea classes remains limited. The Insomnia and Sleep Apnea classes have not yet been successfully identified by the system. This is due to the dominance of data in the None class, causing the model to favor predictions for the majority class.

3.5 Interpretation of Classification Results

The Naïve Bayes algorithm successfully classified the three sleep disorder classes with an accuracy of 71.60%. The best performance was demonstrated in the None class with a precision of 0.72, a recall of 1.00, and an F1-score of 0.83. However, classification capabilities for the Insomnia and Sleep Apnea classes remain limited, particularly for the Sleep Apnea class, which the system has not yet successfully identified. The example of prediction results in Figure 2 demonstrates the system's ability to process complex patient profiles: a 47-year-old female student with good sleep quality (7/10), low stress (4/10), and a Normal BMI yields an 72.08% probability for the No Sleep Disorder class. This result is clinically consistent, as protective factors such as high sleep quality, low stress, and normal body weight are indicators of good sleep health [17].

However, it should be noted that the patient's sleep duration (6.1 hours/day) is still slightly below the American Academy of Sleep Medicine's minimum recommendation of 7 hours per night. This is reflected in the remaining 15% probability for the "Insomnia" class in the prediction results, indicating that the model is capable of capturing the nuances of risk factors proportionally [2].

3.6 Analysis of the Naïve Bayes Algorithm's Performance

The performance of Naïve Bayes in this study aligns with the research by Sari et al. (2025), which compared several machine learning algorithms in the classification of sleep disorders. The results of that study indicate that Naïve Bayes performs reasonably well, though it remains inferior to ensemble algorithms such as Random Forest [11]. This finding is further supported by the research of Anindyar et al. (2025), which demonstrates that the Naïve Bayes algorithm tends to perform better on the majority class compared to the minority class in the classification of sleep disorders [18].

Class distribution imbalance is a major factor affecting model performance, particularly in the Sleep Apnea class, which accounts for only 7.75% of the total dataset. This condition causes the model to be biased toward the majority class (None) and results in a low F1-score for the minority class [13]. This is also consistent with the findings in the Sleep-Disordered Breathing detection system by Ghifari et al. (2023), which used Naïve Bayes with ECG signals and reported an accuracy of 85% under more balanced distribution conditions [17].

3.7 Factors Affecting Accuracy

Analysis of the classification results indicates that sleep quality, sleep duration, and stress levels are the most influential predictors in distinguishing between sleep disorder classes. These findings are consistent with the medical literature, which identifies these three factors as primary determinants of sleep disorders [5]. Additionally, BMI and blood pressure categories significantly contribute to distinguishing cases of Sleep Apnea, which are clinically strongly correlated with obesity and hypertension [2].

Occupational variables also serve as differentiating factors, with office workers tending to have higher stress levels associated with an increased risk of insomnia. Similar findings have been reported by various epidemiological studies showing a significant association between mental workload and insomnia complaints among office workers and students [5], [19].

However, the classification results in this study may still be influenced by the considerable variability in the dataset and the relatively limited amount of data, so the Naïve Bayes model's ability to recognize patterns in each class is not yet fully optimized.

3.8 Comparison with Previous Studies

The 71.60% accuracy achieved in this study indicates that the Naïve Bayes algorithm still possesses sufficient capability for classifying sleep disorders. The study by Prabowo et al. (2020) used Naïve Bayes for classifying sleep disorders at a community health center in Sragen with satisfactory results, although it utilized a different dataset and types of sleep disorders [6]. The results of this study are consistent with the research by Sari et al. (2025), which compared the Decision Tree, Random Forest, and Naïve Bayes algorithms in sleep disorder classification. That study showed that Naïve Bayes is capable of performing classification quite well, although its performance is still lower compared to ensemble-based algorithms such as Random Forest [11].

Although other algorithms such as Random Forest and SVM generally achieve higher accuracy, the advantages of Naïve Bayes lie in:

(1) The interpretability of transparent prediction results through class-specific probability values;
(2) High computational efficiency; and
(3) The ability to perform well with limited-size datasets [9], [10]. These characteristics make Naïve Bayes a suitable candidate for implementation in Clinical Decision Support Systems (CDSS) at healthcare facilities with limited computational resources.

3.9 Implications for the Field of Health Information Management

From a health information management perspective, this study demonstrates the potential of data mining as a tool for the initial screening of sleep disorders that can be integrated into Hospital Information Systems (HIS) or Electronic Health Records (EHR). A Naïve Bayes-based prediction system can provide early warnings to clinicians regarding a patient's risk of sleep disorders based on data routinely collected during visits, without requiring expensive and invasive diagnostic tests such as polysomnography [17].

The results of this study support the urgency of integrating artificial intelligence into clinical workflows to improve the efficiency of sleep disorder detection. By identifying risk factors early, preventive interventions such as sleep hygiene education, stress management, and lifestyle modifications can be provided in a timely manner before the condition progresses into a more serious sleep disorder requiring pharmacological treatment [5], [19].

4. Conclusions

This study successfully applied the Naïve Bayes algorithm to classify sleep disorders using the Sleep Health and Lifestyle Dataset, which consists of 400 data points with 13 attributes. The following are the main conclusions that can be drawn:

1. **Model accuracy:** The Naïve Bayes algorithm achieved an overall accuracy of 71.60% on 81 test data points. These results indicate that while the Naïve Bayes algorithm is capable of classifying sleep health data, the model's performance remains suboptimal due to class distribution imbalance in the dataset.
2. **Performance by class:** The model performed best on the None class with a precision of 0.72, recall of 1.00, and an F1-score of 0.83. Meanwhile, the Insomnia and Sleep Apnea classes were not recognized by the model at all, with precision, recall, and F1-score values of 0.00 for both classes. This condition is directly caused by the dominance of the None class data (72.5%), which causes the model to tend to predict all data into the majority class.
3. **Impact of class imbalance**: Class distribution imbalance is the primary factor limiting the model's performance in this study. The Sleep Apnea class, representing only 7.75%, and the Insomnia class, representing 19.75% of the total data, were insufficiently represented during training, causing the model to fail to learn the patterns distinguishing these two classes from the majority class.
4. **Key predictor factors:** Sleep quality and sleep duration were found to be key predictors in distinguishing between different types of sleep disorders. These findings are consistent with previous research emphasizing the importance of sleep duration for an individual's health and well-being [20].

5. **Implementation potential:** The Naïve Bayes approach has three strategic advantages for implementation in healthcare facilities: interpretability of prediction results through class-specific probability values, high computational efficiency, and the ability to perform optimally on limited-size datasets. This algorithm has the potential to be integrated into Hospital Information Systems (HIS) or Electronic Health Records (EHR) as a Clinical Decision Support System (CDSS) module without requiring expensive diagnostic tests such as polysomnography.

For future research, the following is recommended:

1. Applying class imbalance handling techniques such as SMOTE (Synthetic Minority Over-sampling Technique) or ADASYN as a top priority to improve the model's ability to recognize the Insomnia and Sleep Apnea classes;
2. Using a larger dataset with a more proportional class distribution;
3. comparing the performance of Naïve Bayes with other algorithms such as Random Forest, SVM, and XGBoost on the same dataset;
4. Conducting external validation using clinical data from healthcare facilities in Indonesia; and
5. Considering the addition of biological features such as oximetry data and polysomnography records to improve the representation of minority classes.

Patents

Author Contributions: Conceptualization, O.N.A., A.R.S., A.F.R., and K.F.A.; methodology, O.N.A., K.F.A., and A.R.S.; software, A.R.S. and A.F.R.; validation, A.F.R. and K.F.A.; formal analysis, A.R.S., and A.F.R.; investigation, O.N.A., A.R.S., A.F.R., and K.F.A.; resources, M.Y.; data curation, K.F.A. and A.R.S; writing—original draft preparation, O.N.A.; writing—review and editing, O.N.A. and A.F.R.; visualization, O.N.A., and A.F.R.; supervision, M.Y.; project administration, O.N.A. and K.F.A. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: The authors would like to thank the Health Information Management Study Program of Politeknik Negeri Jember for the academic support provided during this research. The authors also express their sincere *gratitude* to Muhammad Yunus for the guidance, suggestions, and support throughout the research and manuscript preparation process. In addition, the authors appreciate all researchers and authors whose studies were used as references in this journal, as well as all team members for their cooperation and teamwork until this research and article were successfully completed.

References

[1] M. J. Sateia, "International Classification of Sleep Disorders-Third Edition Highlights and Modifications," *Chest*, vol. 146, no. 5, pp. 1387–1394, Nov. 2014, doi: 10.1378/chest.14-0970.

[2] T. Young, P. E. Peppard, and D. J. Gottlieb, "Epidemiology of Obstructive Sleep Apnea: A Population Health Perspective," May 01, 2002. doi: 10.1164/rccm.2109080.

[3] A. V. Benjafield *et al.*, "Estimation of the global prevalence and burden of insomnia: a systematic literature review-based analysis," Aug. 01, 2025, *W.B. Saunders Ltd*. doi: 10.1016/j.smrv.2025.102121.

[4] K. Peltzer and S. Pengpid, "Prevalence, Social and Health Correlates of Insomnia Among Persons 15 Years and Older in Indonesia," *Psychol. Health Med.*, vol. 24, no. 6, pp. 757–768, Jul. 2019, doi: 10.1080/13548506.2019.1566621.

[5] C. M. Morin *et al.*, "Insomnia Disorder," *Nat. Rev. Dis. Primers*, vol. 1, Sep. 2015, doi: 10.1038/nrdp.2015.26.

[6] I. A. Prabowo, D. Remawati, and A. P. W. Wardana, "Klasifikasi Tingkat Gangguan Tidur Menggunakan Algoritma Naïve Bayes," *Jurnal Teknologi Informasi dan Komunikasi (TIKomSiN)*, vol. 8, no. 2, Oct. 2020, doi: 10.30646/tikomsin.v8i2.519.

[7] A. Damayanti and G. Testiana, "Penerapan Data Mining Untuk Prediksi Penyakit Hepatitis C Menggunakan Algoritma Naïve Baye," *Jurnal Manajemen Informatika Jayakarta*, vol. 3, no. 2, pp. 177–186, 2023, doi: 10.52362/jmijayakarta.v3i2.1098.

[8] H. Firda, R. Athiyah, and M. Ihsan Jambak, "Perbandingan Algoritma Klasifikasi Data Mining Model C4.5 Dan Naïve Bayes Untuk Prediksi Keganasan Kanker Payudara," 2024. [Online]. Available: https://jurnal.umj.ac.id/index.php/just-it/index

[9] H. Zhang, "The Optimality of *Naïve*Bayes," 2004. [Online]. Available: www.aaai.org

[10] Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. 2001.

[11] I. Novita Sari, L. Fakih Lidimilah, and A. Lutfi, "Proceeding National Conference Of Research And Community Service Sisi Indonesia 2025 Analisis Perbandingan Algoritma Decision Tree, Random Forest, Dan Naïve Bayes Dalam Klasifikasi Gangguan Tidur," 2025. [Online]. Available: https://conference.sinesia.id/ncrcs-sinesia

[12] Fradillah, M. Fadhiel Alie, and R. Rahmanda, "Model Prediksi Gangguan Tidur berdasarkan Beberapa Faktor menggunakan Machine Learning Prediction of Sleep Disorders based on Several Factors using Machine Learning," 2024. [Online]. Available: www.jurnal.unimed.ac.id

[13] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining Practical Machine Learning Tools and Techniques Fourth Edition*. 2017. [Online]. Available: https://www.elsevier.com

[14] A. Febrisa Sidabutar, R. Habibi, W. Istri, and R. Vokasi, "PERBANDINGAN METODE KLASIFIKASI UNTUK PENGELOMPOKAN RISIKO MAGANG MAHASISWA," 2023.

[15] M. Y. Haffandi, E. Haerani, F. Syafria, and L. Oktavia, "KLASIFIKASI PENYAKIT PARU-PARU DENGAN MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER," *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, vol. 5, no. 2, p. 176, Dec. 2022, doi: 10.37600/tekinkom.v5i2.649.

[16] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.

[17] A. Ghifari, E. R. Widasari, and P. Korespondensi, "Sistem Pendeteksi Sleep-Disordered Breathing Berdasarkan High Dan Low Frequency Menggunakan Metode Naïve Bayes," *Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 2023, doi: 10.25126/jtiik.2023106913.

[18] M. Mega Amelia, B. Maulana Fazrin, Y. Yosua Panjaitan, M. Dicky Kurniawan, and N. Khasanah, "Implementasi Naïve Bayes Untuk Klasifikasi Gangguan Tidur," *Journal Computer Science*, vol. 4, no. 1, 2025.

[19] T. Retno Nurasih, P. Dewi, and I. Heri Susanti, "Hubungan Tingkat Stres Dengan Kualitas Tidur Mahasiswa Yang Sedang Menyusun Tugas Akhir Di Universitas Harapan Bangsa," 2022. [Online]. Available: https://journal-mandiracendikia.com/jik-mc

[20] M. Hirshkowitz *et al.*, "National sleep foundation's sleep time duration recommendations: Methodology and results summary," *Sleep Health*, vol. 1, no. 1, pp. 40–43, Mar. 2015, doi: 10.1016/j.sleh.2014.12.010.