Analysis of Sleep Disorder Classification Using Decision Tree, Naive Bayes, and K-NN

Adithia Gunawan[1], Adriel Walintukan[2], Julio Weol[3], Mutiara Makarawung[4], Cleymency Kaparang[5]

[1,2,3,4,5]Universitas Klabat; Jl. Arnold Mononutu, Airmadidi Bawah, Kec. Airmadidi, Kabupaten Minahasa Utara, Sulawesi Utara, 95371

[1,2,3,4,5]Ilmu Komputer/Sistem Informasi, Fakultas Ilmu Komputer, Universitas Klabat, Airmadidi

e-mail: [1]s22310014@student.unklab.ac.id, [2]s22310041@student.unklab.ac.id,[3]s22219020@student.unklab.ac.id, [4]s22310105@student.unklab.ac.id, [5]s22210615@student.unklab.ac.id

Abstract (Style: Abstract Title)

Sleep disorders are increasingly common health issues affecting all age groups, driven by modern lifestyle imbalances. This study aims to classify sleep disorders using data mining techniques, specifically Decision Tree, Naïve Bayes, and K-Nearest Neighbor (k-NN) algorithms on the Sleep Health and Lifestyle Dataset. The dataset contains various lifestyle and health attributes such as sleep duration, sleep quality, stress levels, physical activity, BMI, and blood pressure. Model performance was evaluated using K-Fold Cross Validation. The results showed that the Decision Tree algorithm achieved the highest accuracy of 88.77%, followed by k-NN and Naïve Bayes with competitive accuracy levels. Moreover, the Decision Tree generated clear and interpretable classification rules, making it suitable for implementation in decision support systems in the healthcare sector. These findings are expected to contribute to the development of AI-based sleep disorder diagnostic systems and promote public awareness of the importance of maintaining a healthy lifestyle.

Keywords— Sleep disorders, data mining, Decision Tree, Naïve Bayes, K-Nearest Neighbor

1. INTRODUCTION

Data mining or data mining in the digital era has become a strategic step for use in various fields. One area that uses data mining is sales and payment transactions [6]. Sleep is one of the basic human needs that plays an important role in maintaining the body's physiological and psychological balance. Good quality sleep contributes significantly to the body's recovery, emotional stability, immune system enhancement, and optimal cognitive performance. Conversely, lack of sleep or prolonged sleep disturbances can have serious adverse effects, such as metabolic disorders, decreased immunity, impaired concentration, and increased risk of various chronic diseases such as hypertension, diabetes mellitus, cardiovascular disorders, and mental health disorders [1], [5].

In the modern era, sleep disorders have become an increasingly common problem across all ages. Factors such as job stress, excessive use of electronic devices, irregular diet, caffeine consumption, lack of physical activity, and general lifestyle imbalance are the main causes of disrupted sleep patterns [2], [5]. Sleep duration is a fundamental indicator of individual health and well-being, and it varies significantly by age group. The National Sleep Foundation recommends 7–9 hours for adults and 7–8 hours for older adults as optimal sleep duration for maintaining health [19]. This condition raises the need for innovative technology-based solutions that can help the public and medical personnel in detecting and classifying sleep disorders quickly, accurately, and efficiently. One of the recent approaches used is the classification of sleep disorders based on EEG signals, as conducted by Utami [16], who employed the Extreme Learning Machine (ELM) algorithm to detect REM Behaviour

Disorder (RBD) and compared it with other algorithms in terms of accuracy and time efficiency. The moment of sleep onset is characterized by gradual changes in many behavioural and physiological characteristics [7].

One promising approach in this context is the use of data mining, particularly classification techniques, to process and analyze the abundance of health data. Data mining allows researchers to explore hidden patterns in large and complex data, and make predictions about certain conditions based on historical data. The goal of classification is to accurately predict the target class for each case in the data [9]. This technique has proven to be effective in various fields, including the healthcare sector [2], [3]. In the context of this research, the Sleep Health and Lifestyle dataset is used, which contains various health and lifestyle attributes such as sleep duration, sleep quality, stress level, physical activity, body mass index (BMI), and blood pressure. This dataset can be a valuable source of information to identify possible sleep disorders in individuals [4].

Several classification algorithms have been widely used for sleep disorder detection and prediction. Among them are Decision Tree, Naïve Bayes (NB), and k-Nearest Neighbor (k-NN). Each algorithm has its own advantages. The Naïve Bayes algorithm, for example, is a simple yet effective Bayes Theorem-based probabilistic classification model, especially when the attributes in the data are mutually independent. Research conducted by Prabowo et al [4] shows that NB is able to classify types of sleep disorders with up to 80% accuracy, with disorder categories such as insomnia, hypersomnia, narcolepsy, sleep terror, and others.

Meanwhile, the Decision Tree algorithm has been proven to excel in producing decision tree-shaped classification models that are not only accurate but also easily understood by end users. Septiani [4] and Nawawi et al. [5] in his research revealed that Decision Tree is able to provide a clear interpretation in mapping factors that affect sleep patterns based on lifestyle attributes, as well as providing tree-based rules that are easily implemented in expert systems.

On the other hand, the k-Nearest Neighbor (k-NN) algorithm is an instance-based classification method that works by calculating the proximity between data. This algorithm is considered suitable for small to medium-sized datasets, and can provide accurate results if the value of the k parameter is determined appropriately. Khasanah et al. [6] shows that k-NN can classify sleep disorders with competitive accuracy, although the performance of this algorithm is greatly affected by data distribution and imbalance between classes.

From the various studies that have been conducted [1], [2], [3], [4], [5], [11] it can be concluded that there is no one algorithm that is absolutely superior in all conditions. Therefore, in this research, a comparative study is conducted on the performance of three classification algorithms, namely Decision Tree, Naïve Bayes, and k-NN, in detecting and classifying sleep disorders on the Sleep Health and Lifestyle Dataset. This study not only aims to find out which algorithm is the most accurate, but also to analyze the effectiveness of each method in processing lifestyle health data, as well as its contribution in supporting artificial intelligence (AI)-based automated diagnosis systems in the health field.

With the results obtained from this research, it is hoped that it can contribute to the development of decision support systems in the health sector, especially in efforts to detect early sleep disorders. In addition, this research is also expected to encourage public awareness of the importance of maintaining sleep patterns and healthy lifestyles through a smart and adaptive technology approach.

2. RESEARCH METHODS

2.1. Data Preparation

The dataset used is the sleep health and lifestyle dataset which we get from the kaggle platform https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset. This data consists of health and lifestyle information related to sleep patterns. Some important variables in this data include Gender, Age, Occupation, Sleep Duration, Quality of Sleep, Physical Activity Level, Stress Level, BMI Category, Blood Pressure, Heart Rate, Daily Steps, Sleep Disorder. This data set consists of 374 observational data with 12 attributes. The performance of data mining algorithms is also influenced by the attributes used in the classification stage, considering that not all data or attributes in the dataset will be used in this process. The purpose of this stage is to ensure that the data used is appropriate to the needs of the analysis [8].

	A	B	C	J	K	L
1	Gender	Age	Occupation	Heart Rate	Daily Steps	Sleep Disorder
2	Male	27	Software Engineer	77	4200	None
3	Male	28	Doctor	75	10000	None
4	Male	28	Doctor	75	10000	None
371	Female	59	Nurse	68	7000	Sleep Apnea
372	Female	59	Nurse	68	7000	Sleep Apnea
373	Female	59	Nurse	68	7000	Sleep Apnea
374	Female	59	Nurse	68	7000	Sleep Apnea
375	Female	59	Nurse	68	7000	Sleep Apnea

Figure 1. Sleep Health and Lifestyle Dataset

After checking the dataset, no missing values or data noise were found, so the data was considered clean and ready for further analysis.

Filter (12 / 12 attributes): Search for Attributes

Name	Type	Missing	Statistics		
Label Sleep Disorder	Nominal	0	Least Insomnia (77)	Most None (219)	Values None (219), Sleep Apnea (78), ...[1 more]
Gender	Nominal	0	Least Female (185)	Most Male (189)	Values Male (189), Female (185)
Age	Integer	0	Min 27	Max 59	Average 42.184
Occupation	Nominal	0	Least Manager (1)	Most Nurse (73)	Values Nurse (73), Doctor (71), ...[9 more]
Sleep Duration	Real	0	Min 5.800	Max 8.500	Average 7.132
Quality of Sleep	Integer	0	Min 4	Max 9	Average 7.313
Physical Activity Level	Integer	0	Min 30	Max 90	Average 59.171
Stress Level	Integer	0	Min 3	Max 8	Average 5.385
BMI Category	Nominal	0	Least Obese (10)	Most Normal (195)	Values Normal (195), Overweight (148), ...[2 more]
Blood Pressure	Nominal	0	Least 122/80 (1)	Most 130/85 (99)	Values 130/85 (99), 125/80 (65), ...[23 more]
Heart Rate	Integer	0	Min 65	Max 86	Average 70.166
Daily Steps	Integer	0	Min 3000	Max 10000	Average 6816.845

Figure 2. no missing values

2.2. *Data Modeling*

2.2.1. Decision Tree Algorithm

Decision Tree Algorithm is a method in machine learning that serves to make decisions or classifications based on a set of rules formed from data. This method has been proven effective in previous studies, such as by Nasrullah [17], who achieved 90% accuracy in classifying product categories using the C4.5 algorithm, showing its reliability for classification tasks involving categorical data. It works by gradually splitting the data, prioritizing the features that most influence the prediction or target category. Each split of the data generates a new branch in the tree structure until it reaches the last node or "leaf", which provides the final decision. The Decision Tree algorithm is an algorithm that can be used to form a decision tree. The C4.5 algorithm is one of the algorithm in decision tree induction, namely ID3 (Iterative Dichotomiser 3) developed by J. Ross Quinlan[13]. Decision tree is one of the methods that is quite easy for humans to interpret. With a tree-like structure, this algorithm is easy to understand even by users who do not have a technical background, as it provides a clear picture of the decision-making process. Decision trees work by changing complex decision making into simple ones so that solutions can be found that are easy to understand and visualize [14]. Decision Tree is widely used in various applications, such as health analysis, credit scoring, and customer segmentation, thanks to its ability to identify complex patterns in data. The algorithm

is suitable for both classification and regression problems, making it very flexible in its application.
Steps taken to get the results using decision tree:
1. Import csv data into Process View by using the Read CSV operator.
2. Define the role using the Set Role operator.
3. Evaluation and validation were tested for accuracy with the K-Fold Cross Validation operator.
4. Connect the process to the Decision Tree operator, Apply Model and Performance (Classification).
5. Process execution.

2.2.2. Naive Bayes Algorithm

Naive Bayes algorithm is a machine learning method used to perform classification based on Bayes' theorem with the assumption of independence between features. Naive Bayes is a classification method in data mining using a simple probabilistic that calculates a set of probabilities by summing the frequencies and combinations of values in a given dataset [10]. This algorithm works by calculating the probability of each class based on the features, then choosing the class with the highest probability as the prediction result. Naive Bayes is often used in the classification of symptoms or clinical conditions due to its ability to efficiently process independent variables. The Naive Bayes method successfully classified depression symptoms based on BDI test results with an accuracy of 91.67%, which is higher than the K-Nearest Neighbor (KNN) method that achieved 86.75% [20]. Naive Bayes belongs to the group of simple yet highly effective classification algorithms, especially for data that has a large number of features. Although its independence assumption is often not fully met in the real world, it still performs well in many real-world cases. One of the advantages of Naive Bayes is its speed in training and prediction. It is suitable for text classification such as spam email filtering, sentiment analysis, and document clustering. Since the process is based on probabilities, the classification results can be easily understood by users.
Steps taken to get the results using decision tree:
1. Import csv data into Process View by using the Read CSV operator.
2. Define the role using the Set Role operator.
3. Evaluation and validation were tested for accuracy with the K-Fold Cross Validation operator.
4. Connect the process to the Naive Bayes operator, Apply Model and Performance (Classification).
5. Process execution.

2.2.3. k-NN(K-Nearest Neighbor) Algorithm

The K-Nearest Neighbor (K-NN) algorithm is a machine learning method used for classification or regression, based on the proximity (distance) of new data to existing data. The main principle of K-NN is that new data will be classified based on the majority of labels from its K nearest neighbors in the feature space. Ensemble-based classification methods, such as bagged trees, have been proven to achieve high accuracy in detecting sleep disorders using only 30-second ECG signals[18]. The K-NN approach is also an example of a lazy learning technique, that is, a technique that waits until the query arrives to generalize beyond the training data [15]. K-NN is a lazy learning algorithm, as the entire prediction process is done when the data is tested. Nonetheless, this algorithm is very effective in various classification problems, especially when the data has a clear pattern based on the distance or similarity between the data. The advantages of K-NN are its simplicity and its ability to capture non-linear relationships in data. However, K-NN can be slow when the data size is very large because it needs to calculate the distance to all data during the prediction process.

This algorithm is widely used in recommendation systems, pattern recognition, and image classification. K-NN is suitable for users who want classification results based on the direct similarity of known data.
Steps taken to get the results using decision tree:

1. Import csv data into Process View by using the Read CSV operator.
2. Define the role using the Set Role operator.
3. Evaluation and validation were tested for accuracy with the K-Fold Cross Validation operator.
4. Connect the process to the k-NN(K-Nearest Neighbor) operator, Apply Model and Performance (Classification).
5. Set the K parameter as required (K = 5, 6, 7, 8).
6. Process execution.

3. RESULT AND DISCUSSION

3.1 *Model Evaluation and Validation Using K-Fold Cross-Validation*

Experiments and model testing were conducted thoroughly with the aim of calculating the accuracy of each algorithm used. In this evaluation process, the cross-validation method was applied as a validation technique to ensure that the results obtained do not depend on a particular data division, but rather represent the general performance of the model against the entire dataset. The use of cross-validation, specifically the k-fold cross-validation technique, allows for a fair comparison between algorithms and provides a more accurate picture of which algorithm has the best predictive ability in classifying sleep disorders based on health and lifestyle data.

3.1.1 Decision Tree

Experiments were conducted using existing operators in rapidminer by testing the K-Fold Cross Validation model for the Decision Tree Algorithm as shown in Figure 3 and Figure 4:

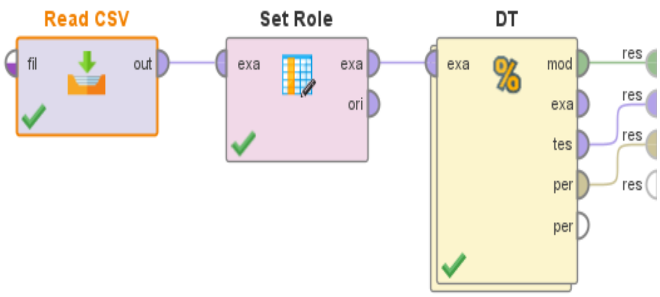

figure 3. Cross Validation Testing

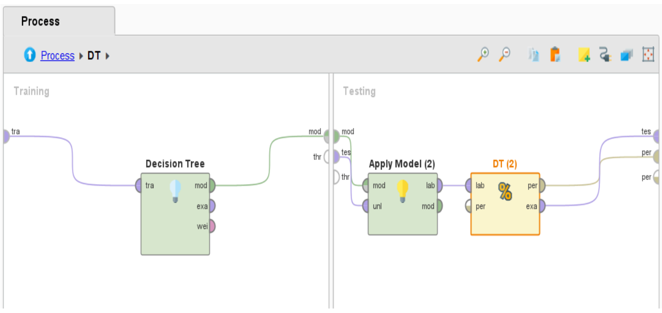

figure 4 Decision Tree Method Testing

3.1.2 k-NN(K-Nearest Neighbor)

Experiments were conducted using operators in rapidminer by testing the K-Fold Cross Validation model for the k-NN (K-Nearest Neighbor) algorithm as shown in Figure 5 and Figure 6:

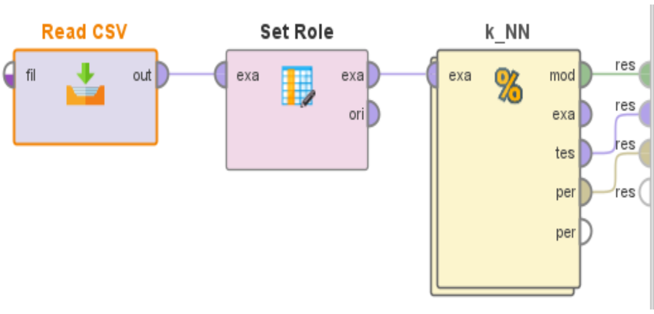

figure 5. Cross Validation Testing

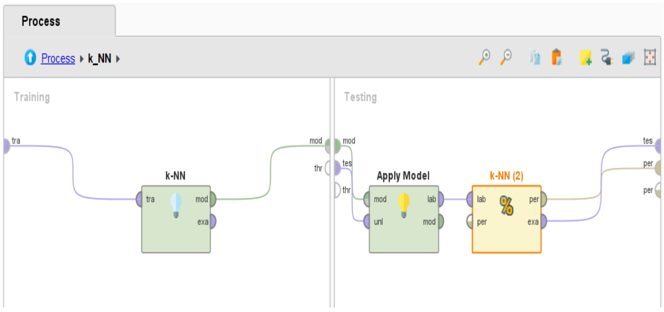

figure 6. k-NN Method Testing

3.1.3 Naive bayes

Experiments were conducted using existing operators in rapidminer by testing the K-Fold Cross Validation model for the Naive Bayes Algorithm as shown in Figure 7 and Figure 8:

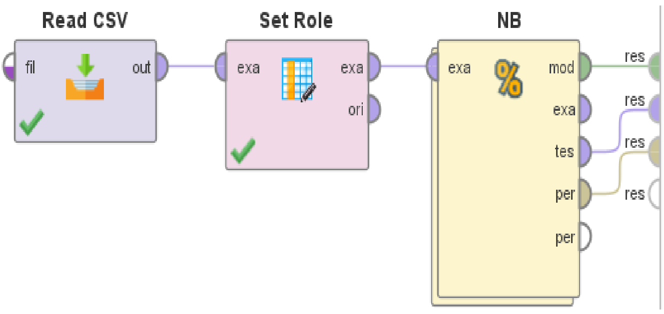

figure 7. Cross Validation Testing

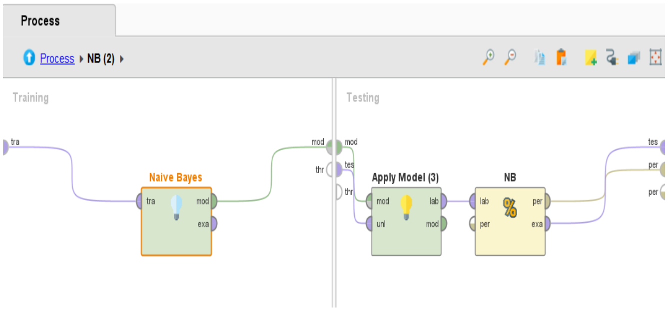

figure 8. Naive Bayes Method Testing

3.2 *Result*

3.2.1 Decision Tree

Table 1. Accuracy of Decision Tree Algorithm

	true None	**true Sleep Apnea**	**true Insomnia**	**class precision**
pred. None	206	7	7	93.61 %
pred. Sleep Apnea	8	64	7	81.01 %
pred. Insomnia	8	7	63	82.89 %
class recall	93.61%	82.05%	81.82%	
Accuray: 88.73% +/- 7.10% (micro average: 88.77%)				

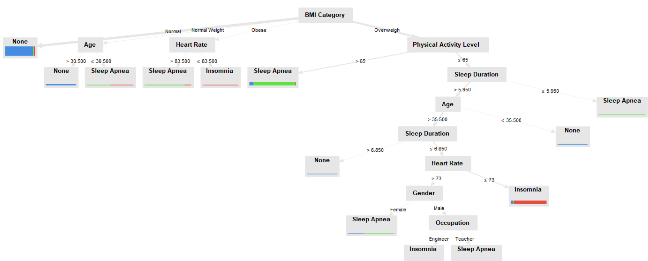

figure 9. Decision Tree Result

Figure 9 shows this experiment generated 13 rules as rules in the classification of sleep disorders. The resulting rules are as follows:

1. If BMI = Normal Weight, then None (No Sleep Disorder)
2. If BMI = Normal Weight and AGE > 30.5 years then None (No Sleep Disorder)
3. If BMI = Normal Weight and AGE <= 30.5 years then Sleep Apnea
4. If BMI = Obese and Heart Rate > 83.5 then Sleep Apnea
5. If BMI = Obese and Heart Rate <= 83, 5 then Insomnia
6. If BMI = Overweight and Physical Activity Level > 65 then Sleep Apnea
7. If BMI = Overweight and Physical Activity Level > 47.5 and <= 65 then None (No Sleep Disorder)
8. If BMI = Overweight and Physical Activity Level <= 47.5 and Sleep Duration > 5.95

and Age > 42.5 then Insomnia.

9. If BMI = Overweight and Physical Activity Level <= 47,5 dan Sleep Duration > 5.95 and Age <= 42.5 and Gender = Female, then Sleep Apnea.
10. If BMI = Overweight and Physical Activity Level <= 47,5 dan Sleep Duration > 5.95 and Age <= 37.5 and Gender = Male, then Insomnia.
11. If BMI = Overweight and Physical Activity Level <= 47,5 dan Sleep Duration > 5.95 and Age <= 37.5 and Gender = Male, then Sleep Apnea.
12. If BMI = Overweight and Physical Activity Level <= 47.5 and Sleep Duration <= 5.95 and Age <= 35.5 then None (No Sleep Disorder).
13. If BMI = Overweight and Physical Activity Level <= 47.5 and Sleep Duration <= 5.95 then Sleep Apnea.

3.2.2 Naive Bayes

Table 2. Accuracy of Naive Bayes

	true None	**true Sleep Apnea**	**true Insomnia**	**class precision**
pred. None	206	5	7	94.50 %
pred. Sleep Apnea	5	67	5	87.01 %
pred. Insomnia	8	6	65	82.28 %
class recall	94.06%	85.90%	84.42%	
Accuray: 90.36% +/- 5.12% (micro average: 90.37%)				

SimpleDistribution

```
Distribution model for label attribute Sleep Disorder


Class None (0.586)
11 distributions

Class Sleep Apnea (0.209)
11 distributions

Class Insomnia (0.206)
11 distributions
```

figure 10. Naive Bayes Result

Figure 10 shows the distribution of the data shows that the majority of the data in this dataset belongs to the None (no sleep disorder) class at 58.6%. Meanwhile, the other two sleep disorder classes, namely:Sleep Apnea has a proportion of about 20.9%.Insomnia has a proportion of about 20.6%.The same number of distributions (11 distributions) for these three classes indicates that the data has been analyzed by proportion per attribute, but in terms of quantity, the None class dominates.

3.2.3 k-NN

Table 3. Accuracy of K-Nearest Neighbor (K-NN) Algorithm Table Experimental results by changing the value of K(5, 6, 7, 8)

K value	Accuracy Result
5	85.56%
6	84.76%
7	86.10%
8	86.10%

KNNClassification

```
Weighted 5-Nearest Neighbour model for classification.
The model contains 374 examples with 11 dimensions of the following classes:
  None
  Sleep Apnea
  Insomnia
```

figure 11. k-NN result

Figure 11 shows that a classification model has been created using the K-Nearest Neighbors (KNN) algorithm, to be precise:

1. It uses 5 nearest neighbors and a weighted approach in making classification decisions.
2. The model was built from 374 examples.
3. The goal of the model is to classify the data into three classes, namely:
    1. None (no sleep disorder),
    2. Sleep Apnea (breathing disorder during sleep),
    3. Insomnia (sleep difficulty).

4. CONCLUSION

The conclusion of this research shows that the Naïve Bayes algorithm provides the best performance in sleep disorder classification, with the highest accuracy rate reaching 90.37%. The main advantage of this algorithm lies in its prediction speed and efficiency in handling large datasets with relatively simple attribute distributions. Meanwhile, the Decision Tree algorithm also showed competitive performance, with the main advantage being its ability to generate classification rules that can be easily interpreted. This makes Decision Tree very feasible to be implemented in decision support systems in the health sector, especially to assist medical personnel in diagnosing sleep disorders based on lifestyle and health parameters.

In addition, the k-Nearest Neighbor (k-NN) algorithm also provides quite competitive results compared to the other two algorithms. The advantage of k-NN lies in its flexibility in measuring proximity between data, although its performance is strongly influenced by the distribution of data and the selection of parameter k. Based on the resulting Decision Tree model, the factors that have the most influence on sleep disturbance include BMI, age, heart rate, physical activity level, sleep duration, and gender. These findings confirm the importance of lifestyle and physiological health factors in determining the risk of sleep disorders.

The distribution of data in the dataset also shows inequality between classes, where the “None” category (no sleep disorder) dominates as much as 58.6% of the total data. Meanwhile, the other two categories, “Sleep Apnea” and “Insomnia”, are around 20% each, indicating an imbalance in the data that affects the classification results. Therefore, further research is recommended to consider data balancing techniques such as oversampling or undersampling to improve classification performance in minority classes.

REFERENCES

[1] I. A. Prabowo, D. Remawati, dan A. P. W. Wardana, “Klasifikasi Tingkat Gangguan Tidur Menggunakan Algoritma Naïve Bayes,” *Jurnal TIKomSiN*, vol. 8, no. 2, pp. 40–47, Okt. 2020.

[2] Y. Mardi, “Data Mining: Klasifikasi Menggunakan Algoritma C4.5,” *Jurnal Edik Informatika*, vol. 2, no. 2, pp. 213–219, 2016.

[3] I. Nawawi dan Z. Fatah, “Penerapan Decision Trees dalam Mendeteksi Pola Tidur Sehat Berdasarkan Kebiasaan Gaya Hidup,” *Jurnal Ilmiah Sains Teknologi dan Informasi*, vol. 2, no. 4, pp. 34–41, Okt. 2024.

[4] W. D. Septiani, “Klasifikasi Gangguan Tidur Menggunakan Metode Decision Tree dan Algoritma Genetika,” *Jurnal Minfo Polgan*, vol. 13, no. 2, pp. 2668–2671, Feb. 2025.

[5] N. Khasanah, D. U. E. Saputri, F. Aziz, dan T. Hidayat, “Studi Perbandingan Algoritma Random Forest dan K-Nearest Neighbors (KNN) dalam Klasifikasi Gangguan Tidur,” *Computer Science (CO-SCIENCE)*, vol. 5, no. 1, pp. 17–20, Jan. 2025.

[6] B. S. B. Sembiring, M. Pandia, H. Sembiring, dan D. Margaretta, “Naïve Bayes Classifier and Decision Tree Algorithms for Classifying Payment Data,” *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 1, pp. 592–600, Aug. 2023.

[7] S. Chokroverty, "Overview of sleep & sleep disorders," *Indian Journal of Medical Research*, vol. 131, pp. 126–140, Feb. 2010.

[8] M. Kamil dan Y. Erliani, "Implementasi Algoritma C4.5 Untuk Memprediksi Penjualan Sepatu Boots Terlaris di Toko Parabellum," *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 9, no. 3, pp. 852–862, Jul.–Sep. 2025.

[9] V. Krishnaiah, G. Narsimha, and N. S. Chandra, "Survey of classification techniques in data mining," *International Journal of Computer Sciences and Engineering*, vol. 2, no. 9, pp. 65–74, Sep. 2014.

[10] G. D. Rahayuningtyas, "Prediksi Sleep Disorder Menggunakan Metode Algoritma Naïve Bayes Gaussian," *SABER: Jurnal Teknik Informatika, Sains dan Ilmu Komunikasi*, vol. 3, no. 1, pp. 283–291, Jan. 2025.

[11] A. S. Nugroho, D. I. Y. Yuniarti, and E. D. Purnama, "Klasifikasi Sampah Organik dan Anorganik Menggunakan Convolutional Neural Network (CNN)," *Jurnal Teknik ITS*, vol. 10, no. 2, pp. A121–A125, 2021.

[12] A. A. Haris, M. A. Suryani, and A. R. Widodo, "Perbandingan Algoritma Naive Bayes dan C4.5 dalam Menentukan Kelulusan Mahasiswa," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 6, no. 5, pp. 1025–1032, Oct. 2019.

[13] Y. I. Kurniawan, “Perbandingan Algoritma Naive Bayes dan C.45 dalam Klasifikasi Data Mining,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 4, pp. 455–464, Oct. 2018.

[14] H. Lestiawan, C. Jatmoko, F. Agustina, D. Sinaga, and L. Erawan, “Prediction of Sleep Disorders Based on Occupation and Lifestyle: Performance Comparison of Decision Tree, Random Forest, and Naïve Bayes Classifier,” 2023.

[15] Xindong Wu, Vipin Kumar - The Top Ten Algorithms in Data Mining - 2009.pdf_buku13.

[16] A. N. Utami, "Klasifikasi Gangguan Tidur REM Behaviour Disorder Berdasarkan Sinyal EEG Menggunakan Machine Learning," *Jurnal Sistem Cerdas*, vol. 3, no. 2, pp. 216–230, 2020.

[17] A. H. Nasrullah, "Implementasi Algoritma Decision Tree untuk Klasifikasi Produk Laris," *Jurnal Ilmiah Ilmu Komputer*, vol. 7, no. 2, pp. 45–51, Sep. 2021.

[18] E. R. Widasari, K. Tanno, dan H. Tamura, "Automatic Sleep Disorders Classification Using Ensemble of Bagged Tree Based on Sleep Quality Features," *Electronics*, vol. 9, no. 3, p. 512, Mar. 2020.

[19] 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.

[20] N. R. Setiawan, "Perbandingan Klasifikasi Gejala Depresi Menggunakan K-Nearest Neighbor dan Naive Bayesian Classification," *Jurnal Teknik Elektro dan Teknologi Informasi, Universitas Gadjah Mada*, 2014.