JURNAL RISET INFORMATIKA **P-ISSN: 2656-1743 |E-ISSN: 2656-1735**
Vol. 8, No. 1. December 2025 DOI: https://doi.org/10.34288/jri.v8i1.456
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

ENHANCING SLEEP QUALITY PREDICTION THROUGH SMOTE-BASED DATA BALANCING AND HYBRID MACHINE LEARNING MODELS

Ami Rahmawati[-1], Ita Yulianti[-2], Ani Oktarini Sari[3], Siti Nurajizah[-4] , Hikmatulloh[-5]

Faculty of Information Technology [-1, -3, -5]
Universitas Nusa Mandiri
ami.amv@nusamandiri.ac.id[-1], ani.aos@nusamandiri.ac.id[-3], hikmatulloh.hkl@nusamandiri.ac.id[-5]

Faculty of Engineering and Informatics[-2, -4]
Universitas Bina Sarana Informatika
ita.iyi@bsi.ac.id[-2], siti.snz@bsi.ac.id[-4]

Abstract

Sleep is a vital aspect in maintaining a person's physical and psychological balance. Poor sleep quality can reduce physical and cognitive performance, increasing the risk of various health problems. This study aims to develop a predictive model for sleep quality based on factors such as lifestyle, stress, daily activities, and caffeine consumption, using XGBoost combined with Recursive Feature Elimination (RFE). XGBoost was chosen for its ability to handle imbalanced datasets and heterogeneous features, while RFE helps simplify the model without losing important information. In the data pre-processing stage, a class imbalance was found, so the Synthetic Minority Over-sampling Technique (SMOTE) process was carried out to balance the proportion of the minority class. The dataset in this study was divided into two parts, namely 80% as training data and 20% as testing data, and validated using cross-validation to ensure generalization. The results show very high model performance with an accuracy of 99.79% on training data, 99.63% on cross-validation, and 99.10% on testing data. This model was then developed into a web application for practical use in analyzing sleep quality prediction. This study emphasizes the methodological contribution of a SMOTE-based hybrid machine learning model and its ready-to-use application implementation, while also opening opportunities for further testing on more diverse datasets and evaluating biases caused by synthetic data.

Keywords: Prediction; RFE; Sleep Quality; SMOTE; XGBoost

Abstrak

Tidur merupakan aspek penting dalam menjaga keseimbangan fisik dan psikologis seseorang. Kualitas tidur yang buruk dapat menurunkan kinerja fisik dan kognitif, serta meningkatkan risiko berbagai masalah kesehatan. Penelitian ini bertujuan untuk mengembangkan model prediktif untuk kualitas tidur berdasarkan faktor-faktor seperti gaya hidup, stres, aktivitas sehari-hari, dan konsumsi kafein, menggunakan XGBoost yang dikombinasikan dengan Recursive Feature Elimination (RFE). XGBoost dipilih karena kemampuannya untuk menangani dataset yang tidak seimbang dan fitur yang heterogen, sementara RFE membantu menyederhanakan model tanpa kehilangan informasi penting. Pada tahap pra-pemrosesan data, ditemukan ketidakseimbangan kelas, sehingga dilakukan proses Synthetic Minority Over-sampling Technique (SMOTE) untuk menyeimbangkan proporsi kelas minoritas. Dataset pada penelitian ini dibagi menjadi dua bagian, yaitu 80% sebagai data training dan 20% sebagai data testing, dan divalidasi menggunakan cross-validation untuk memastikan generalisasi. Hasilnya menunjukkan performa model yang sangat tinggi dengan akurasi 99.79% pada data training, 99.63% pada cross-validation, dan 99.10% pada data testing. Model ini kemudian dikembangkan menjadi aplikasi web untuk penggunaan praktis dalam menganalisis prediksi kualitas tidur. Penelitian ini menekankan kontribusi metodologis dari model pembelajaran mesin hibrida berbasis SMOTE dan implementasi aplikasinya yang siap pakai, sekaligus membuka peluang untuk pengujian lebih lanjut pada dataset yang lebih beragam dan mengevaluasi bias yang disebabkan oleh data sintetis.

Kata kunci: Prediksi, RFE; Kualitas Tidur; SMOTE; XGBoost

INTRODUCTION

Sleep is one of the vital aspects in maintaining a person's physical and psychological balance (Nugraha et al., 2025). Suboptimal sleep quality can reduce physical and cognitive performance, and increase the risk of various health problems such as obesity, hypertension, cardiovascular disorders, and mental disorders. (Putra & Hidayat, 2024). Modern lifestyles today often demand that individuals work and be active beyond normal working hours (Thanri et al., 2025) which ultimately leads to a shift in sleep patterns characterized by reduced duration and decreased sleep efficiency (Nugraha et al., 2025). Beside these conditions, sleep quality is also influenced by various other factors, including caffeine consumption, environmental conditions, stress levels, lifestyle, and overall health (Permata et al., 2023).

One form of caffeine consumption behavior that is widely prevalent in society today is the habit of drinking coffee (Tresnawulan et al., 2024). Coffee is often consumed to restore alertness, boost energy, and eliminate drowsiness (Meiranny & Chabibah, 2022). The majority of people today think that coffee consumption is not just for energy, but has also become a social habit (Agustiana P & Nafisah, 2024). Excessive coffee consumption can spur the heart to beat faster, making it difficult to fall asleep (Ranti et al., 2022). This duality illustrates that coffee consumption can have multiple health effects depending on the timing and amount of consumption (Mansur et al., 2025).

Differences in metabolism, psychological state, and lifestyle mean that individuals do not always show the same response to coffee consumption (Priambodo & Chozanah, 2020). This complexity makes it necessary to comprehensively consider a wide range of variables in analyzing sleep quality, rather than focusing on a single factor. The importance of knowing what potential risk factors are associated with sleep quality can minimize health risks (Henrich et al., 2023) and become the basis for building data science-based prediction models. Through this research, data is systematically processed by integrating statistical techniques, machine learning, and data management, so as to extract patterns from the data (Sastra & Sabri, 2025) and produce more accurate predictions of sleep quality.

There are several studies that discuss sleep quality. For example, research from (Nawawi & Fatah, 2024) which analyzes the factors that affect sleep quality using the Decision Tree algorithm. The results of this study focus more on data exploration to see the most influential factors and display the tree structure of the model built without developing it into a system. Furthermore, research from (D. Sari, 2024) which applied SVM and Neural Network algorithms to build a model to predict sleep disorders in Sleep Health and Lifestyle data. In this study, sleep quality was classified into three classes, including insomnia, none, and sleep apnea. The results showed that both algorithms used were able to provide good classification performance with an SVM accuracy value of 90.1% and NN of 91.2%. Other research conducted by (Khasanah et al., 2025) which compares the performance between Random Forest and KNN algorithms in the classification of sleep disorders. The results stated that the accuracy rate of the Random Forest algorithm was able to reach 89.69% and showed a more stable performance when compared to KNN.

Referring to this review, this study aims to build a model that can predict sleep quality using a more efficient and accurate approach through the implementation of the XGBoost algorithm combined with the Recursive Feature Elimination (RFE) method. The XGBoost algorithm was chosen because it has a regularization function that can reduce the occurrence of overfitting and is relatively more robust in handling outliers to predict sleep quality using a more efficient and accurate approach through the implementation of the XGBoost algorithm combined with the Recursive Feature Elimination (RFE) method. The XGBoost algorithm was chosen because it has a regularization function that can reduce overfitting and is relatively more robust in handling outliers (Mujiyono et al., 2025). The XGBoost algorithm uses a decision tree as a base model (Mawardi et al., 2024) by building predictions through a boosting technique that is better than traditional gradient boosting, because it is able to control model complexity and optimize the computational process (Firdaus et al., 2025) so that the system can be built more quickly, efficiently, and accurately (A. P. Sari et al., 2024). Although XGBoost has a robust regularization mechanism, the quality of the resulting prediction is still prone to bias if the class distribution in the dataset is unbalanced. Therefore, to overcome this problem, this research applies SMOTE (Synthetic Minority Oversampling Technique) which serves to balance the class distribution by generating synthetic samples in the minority class. This approach is expected to improve the XGBoost model to learn patterns more

proportionally so as to maximize the sensitivity and accuracy of predictions on minority classes. In addition, the modeling process is also equipped with RFE (Recursive Feature Elimination). Where, RFE is used to perform feature selection by reducing the data dimension through a gradual removal process of less relevant features (Tinaliah & Elizabeth, 2024). This step is done so that XGBoost can focus on variables that have a significant contribution, so that the prediction performance becomes more stable and efficient.

In contrast to most previous studies that focus on predicting sleep quality from the perspective of specific disorders, this study focuses on predicting sleep quality in general by utilizing simple data from daily habits, such as coffee consumption, sleep duration, physical activity, and other lifestyles. This research integrates XGBoost and RFE in the modeling process to improve the efficiency and stability of the prediction. In addition, this research also uses the SMOTE technique to overcome class imbalance in the data. The entire set of data processing to modeling in this research is implemented into a web-based programming that is intended to be easily accessible to the public, thus helping users understand how their daily habits affect sleep quality.

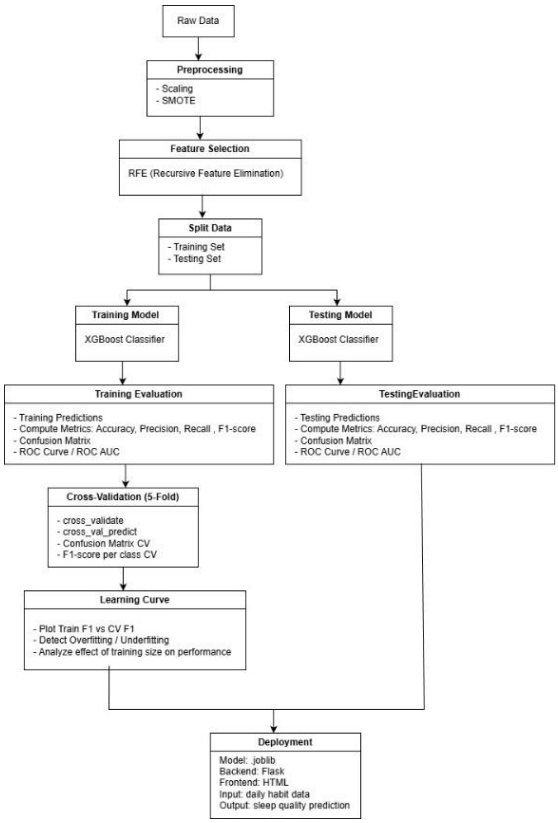

Figure 1. XGBoost Model Evaluation Pipeline with RFE & SMOTE

RESEARCH METHODS

In order for the model in this study to be built and evaluated systematically, a pipeline was created that describes the workflow from the initial stage to the final result. This pipeline aims to ensure that each step can be monitored, optimized, and executed consistently.

The figure above illustrates the entire process flow carried out to obtain the results of the prediction analysis in this study. The research process begins with determining the topic and selecting the dataset to be processed, namely data related to sleep quality. The Global Coffee & Health dataset is used in this study for the classification of four sleep quality classes (excellent, good, fair, poor). The preprocessing stage begins by performing a one-hot encoding technique on relevant categorical variables according to the needs of the model, then continued with feature scaling to normalize the scale of the data, and applying the SMOTE (Synthetic Minority Oversampling Technique) technique with a random_state = 42 setting to overcome the imbalance in the target prediction class. After the preprocessing stage is complete, through the feature selection process, the dataset is then selected using the RFE (Recursive Feature Elimination) technique by selecting the 10 most relevant features, so that the model is more efficient and focused on variables that make a significant contribution to sleep quality prediction. Furthermore, the dataset is then divided into two

parts consisting of training set and testing set with a ratio of 80:20. This proportion is used to provide a large enough portion of training data so that the model can learn patterns and data structures more comprehensively. With the majority of the training data, the model has a wider chance of recognizing class variations in the dataset. Meanwhile, the remaining 20% of the data was used as a test set to ensure that the performance evaluation was performed on data that was completely unseen during the training process.

The main model used is XGBoost with logloss evaluation function, which is suitable for multiclass classification and provides stability to the probability estimation. Then, performance evaluation is carried out through prediction and calculation of metrics such as accuracy, precision, recall, F1-score, confusion matrix, and ROC curve / ROC AUC. To ensure the stability and generalizability of the resulting model, a 5-fold cross-validation was also performed, which resulted in the average metrics with standard deviation, per-sample prediction, confusion matrix, and F1-score per class. The use of 5-fold was chosen because it offers more stable performance estimation with faster computation time, reduces evaluation bias, and ensures that the results are not dependent on one particular data division. Finally, the learning curve was analyzed to compare the training and cross-validation performance, detect the possibility of overfitting or underfitting, and see the effect of training size on model performance. In addition to the main evaluation, the learning curve is used to see if the model performance changes as the training data increases and to identify potential overfitting or underfitting.

After the model is declared stable, the deployment stage is performed. The trained model is saved in .joblib format for easy reloading without retraining. The model is then integrated into a web application built using Flask as the backend. The application accepts input in the form of the same parameters as the selected features from the RFE process, then interprets them using a pipeline identical to the training process, and produces sleep quality predictions.

Types of research

This research belongs to the type of quantitative research using a predictive analysis approach. The main objective of this research is to build a model that is able to predict a person's sleep quality based on lifestyle factors and coffee consumption by utilizing secondary data from Kaggle. This research is a type of quantitative research using a predictive analysis approach. The main objective of this research is to build a model that is able to predict a person's sleep quality based on lifestyle factors and coffee consumption by utilizing secondary data from Kaggle. In addition to data analysis, this research also developed the resulting model into a web-based application using the Python programming language specifically designed to accept input in the form of lifestyle variables (such as frequency of coffee drinking, stress level, and physical activity) and provide output in the form of sleep quality prediction. Thus, this research is not only analytical but also applicative, because the resulting modeling is implemented into the form of a system.

Research Target / Subject

The subject of this study was individual health and lifestyle data publicly obtained through the Kaggle Repository. The data is a Global Coffee Health dataset that contains information including daily coffee consumption intake, physical activity, stress levels, and other health indicators, including sleep duration and quality. Therefore, the target of this research is to produce a prediction model that is able to accurately estimate a person's sleep quality level based on the analysis of the patterns and relationships between variables contained in the dataset.

Data Collection

In this research, data collection is done by sourcing secondary data that is publicly available through the Kaggle Repository platform. As for the dataset used, the Global Coffee Health Dataset is provided in CSV format with several numerical and categorical attributes. This dataset contains around 10,000 synthetic records that describe coffee consumption patterns, sleep behavior, and health conditions of individuals from various countries. The main focus of this research lies on the sleep quality variable which is the target of prediction. Meanwhile, other variables are used as determining factors in building a model that is able to predict sleep quality. This research dataset can be downloaded through the following official link : https://www.kaggle.com/datasets/uom190346a/global-coffee-health-dataset/data

RESULTS AND DISCUSSION

Dataset

The data in this study was obtained from the Kaggle platform and consists of 15 main

attributes that describe an individual's overall profile, so that it can be utilized to predict sleep quality. These attributes include Gender, Country, Coffee Intake, Caffeine, Sleep Hours, BMI, Heart Rate, Stress Level, Physical Activity Hours, Health Issues, Occupation, Smoking, Alcohol Consumption, and Sleep Quality.

Preprocessing

a. Imputation Missing Value

In the sleep quality prediction dataset, an empty value was found in the Health Issues attribute. To overcome this, an imputation process was carried out by replacing the missing values using the 'Unknown' category. This approach is not just filling in the blanks, but also maintaining the completeness of the dataset so that the machine learning algorithm can utilize the entire sample without losing information. The results of the imputation process are presented in Figure 2.

```
Health_Issues
Unknown     5941
Mild        3579
Moderate     463
Severe        17
Name: count, dtype: int64
```

Figure 2. Imputation Missing Value

b. Encoding Categorical Data

The attributes in this research dataset consist of two main data types, namely categorical and numerical. Therefore, the categorical attributes need to be converted to numeric in order to be processed computationally. This process is done by applying the Label Encoding and One-Hot Encoding methods. Label Encoding preserves the order or level of the relevant categories, while One-Hot Encoding prevents the model from assuming ordinal relationships in nominal categories. This encoding process ensures the dataset is ready to be used for model training by retaining the structural information of the categorical attributes. Figure 3 displays the conversion results of the process.

	ID	Age	Gender	Coffee_Intake	Caffeine_mg	Sleep_Hours	BMI	Heart_Rate	Stress_Level	Physical_Activity_Hours	...	Country_Japan
0	1	40	1	35	3281	75	249	78	0	145	...	0
1	2	33	1	10	941	62	200	67	0	110	...	0
2	3	42	1	53	5037	59	227	59	1	112	...	0
3	4	53	1	26	2492	73	247	71	0	66	...	0
4	5	32	0	31	2980	53	241	76	1	85	...	0
...	...	...	...	...	...	...	...	...	...	...	...	...
9995	9996	50	0	21	1968	60	305	50	1	101	...	1
9996	9997	18	0	34	3192	58	191	71	1	116	...	0
9997	9998	26	1	16	1534	71	251	66	0	137	...	0
9998	9999	40	0	34	3271	70	193	80	0	1	...	0
9999	10000	42	0	29	2775	64	281	72	0	98	...	0

10000 rows × 34 columns

Figure 3. Categorical Data Coding

c. Over-Sampling SMOTE

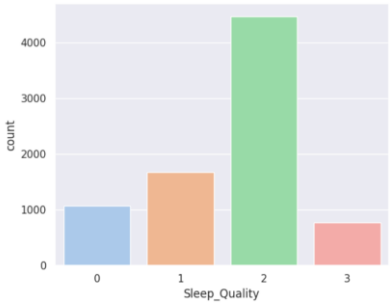

Figure 4. Sleep Quality Class

Based on Figure 4, the data distribution in each class shows a class imbalance, where the "Good" class dominates the number of samples compared to other classes. This imbalance can lead to prediction bias towards the majority class if not addressed. To address this, the Synthetic Minority Oversampling Technique (SMOTE) was applied, which adds synthetic samples to the minority class without subtracting from the original data, so that each class has a balanced amount. After applying SMOTE, the Sleep_Quality attribute shows a balanced distribution, with Excellent, Fair, Good, and Poor classes having 4,475 samples each. This approach enhances the model to recognize balanced data patterns from all classes, improves generalization ability, and minimizes bias towards the majority class. A visualization of the SMOTE results is shown in Figure 5.

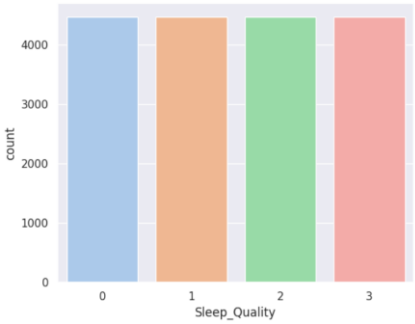

Figure 5. SMOTE Implementation Result

Feature Selection

The next stage is feature selection, which aims to identify the most relevant attributes that affect sleep quality.

P-ISSN: 2656-1743 | E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v8i1.456
JURNAL RISET INFORMATIKA
Vol. 8, No. 1. December 2025
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

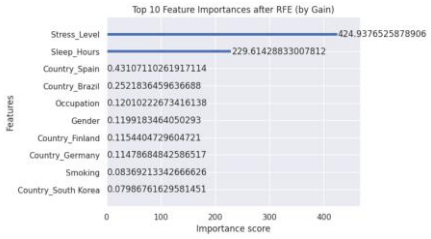

Figure 6. Feature Selection Results

At this stage, information is obtained about the ten most important features that are retained after the selection process using the Recursive Feature Elimination (RFE) method (See Figure 6). The importance of each feature is measured based on the Gain value, which is the amount of contribution of the feature to the improvement of the model performance when used in the data separation process on the decision tree. The Stress_Level feature has the highest gain value of 424.94, indicating that this variable has the most influence on the model's prediction results. This is followed by the Sleep_Hours feature which occupies the second position with a gain value of 229.61, signifying its significant role in determining sleep quality. Meanwhile, the Country_Spain feature has a gain value of 0.43, which indicates a small influence on the prediction results. Meanwhile, for other features such as Country_Brazil, Occupation, Gender, Country_Finland, Country_Germany, Smoking, and Country_South Korea have gain values ranging from 0.25 to 0.08, so their contribution to the model is relatively low.

Modelling Results

In the process of modeling to predict sleep quality, the dataset is divided into two parts, 80% as training data and 20% as testing data. In addition, cross-validation is performed on the training data to ensure more stable model performance and avoid overfitting. The classification process is performed using the RFE algorithm. After the model is formed, performance evaluation is carried out using several metrics, including Accuracy, ROC, Precision, Recall, and F1-Score to assess the overall performance of the model.

a. Evaluation on Training Data

Table 1. Training Data Results

No.	Matrik Evaluasi	Score
1	Accuracy	99,79%
2	Precision	99,79%
3	Recall	99,79%
4	F1 Score	99,79%

Based on the results shown in Table 1, the model at the training stage achieved an accuracy level of 0.9979 or 99.79%, with relatively similar precision, recall, and F1-score values. The high precision and recall values (≈0.9979) indicate that the model is able to identify positive classes with a very low error rate, both in the form of false positives and false negatives.

The existence of similar values in all evaluation metrics (Accuracy, Precision, Recall, and F1-score) is a relatively rare condition. This occurs when the number of prediction errors is very small and proportional between false positives and false negatives. Overall, these results reflect that the model has a very effective learning ability towards the data patterns in the training stage. The Confusion Matrix of the tested model is presented below. This matrix shows the distribution of correct and incorrect predictions, thus showing how Precision, Recall, F1-score, and Accuracy can reach the same value.

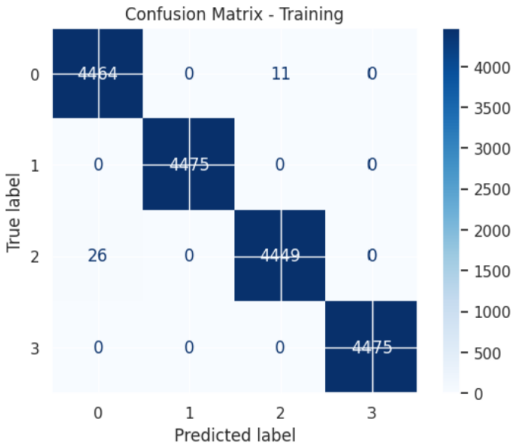

Figure 7. Confusion Matrix Image

The Confusion Matrix in Figure 7 illustrates the classification performance of the model on the training data. For the Excellent class, out of a total of 4,475 data, 4,464 samples were correctly classified, while 11 samples were mistakenly identified as the Good class. The Fair class showed perfect results with 100% accuracy, where all 4,475 data were correctly classified. In the Good class, a total of 4,449 data were correctly classified, while 26 data were incorrectly categorized as the Excellent class. The Poor class also achieved perfect accuracy with all 4,475 data correctly predicted. These results show that the model has a very high level of accuracy on the training data, with very small classification errors, especially between the Excellent and Good classes. This finding is in line

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v8i1.456
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

with the results of the previous evaluation metrics, which confirmed the excellent classification performance at the training stage.

b. Evaluation on Cross Validation

Table 2. Cross Validation Results

No.	Matrik Evaluasi	Score
1	Accuracy	99,63%
2	Precision	99,63%
3	Recall	99,63%
4	F1 Score	99,63%

The cross-validation results shown in Table 2 using the 5-Fold method show that the model has a very consistent and high level of performance across data shares. Accuracy, Precision, Recall, and F1-score values each reached 0.9963 with a very small standard deviation (±0.0015). This achievement shows that the model is able to maintain an average accuracy of 99.63% at each fold, with almost insignificant variations in performance between folds. This finding indicates that the model has excellent stability, not only excelling at one particular subset of data, but also consistently providing accurate prediction results across multiple data divisions during the validation process. Figure 8 displays the Confusion Matrix illustrating the results of the model evaluation at the cross-validation stage.

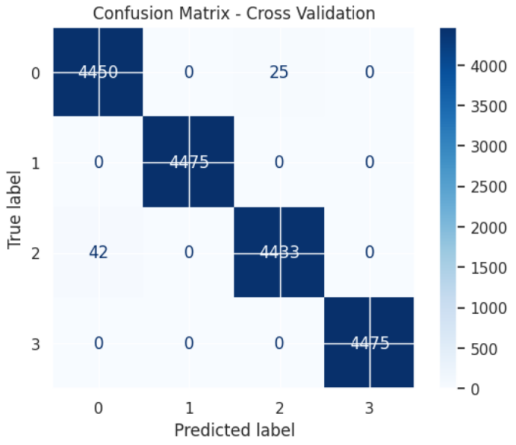

Figure 8. Confusion Matrix Cross Validation

Figure 8 shows the Confusion Matrix of the model evaluation results at the cross-validation stage. For the Excellent class, out of a total of 4,475 data, 4,450 samples were correctly classified, while the other 25 samples were mistakenly classified as the Good class. The Fair class showed perfect results with all 4,475 data correctly classified without error. In the Good class, a total of 4,433 data were correctly classified, while 42 data were incorrectly identified as the Excellent class. Meanwhile, the Poor class also achieved perfect accuracy, with all 4,475 data correctly predicted. These results show that the model has excellent classification capabilities at the cross-validation stage, with minimal prediction errors, especially between the Excellent and Good classes.

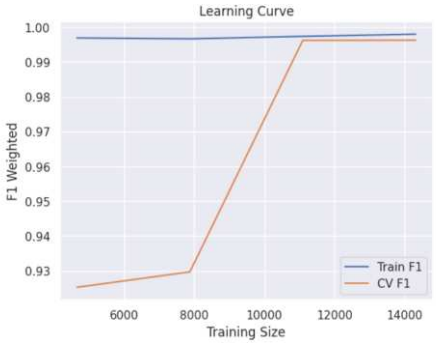

Figure 9. Learning Curve

Figure 9 presents the results of the learning curve based on the weighted F1-score. On small data (5,000-8,000 samples), cross-validation F1 (≈0.93) is lower than train F1 (≈0.998), indicating overfitting and suboptimal model generalization. With the increase of data up to 12,000-15,000 samples, the cross-validation F1 increased almost to the level of train F1 (≈0.999), indicating an improvement in the generalization ability and stability of the model. This trend confirms that the additional training data is effective in reducing overfitting and strengthening pattern representation, while the model performance is close to the optimal limit, making further improvement marginal.

c. Evaluation on Testing Data

Table 3. Testing Data Results

No.	Matrik Evaluasi	Score
1	Accuracy	99,10%
2	Precision	99,10%
3	Recall	99,10%
4	F1 Score	99,10%

Based on the results shown in Table 3, the model performs very well with an accuracy of 0.991, as well as precision, recall, and F1-score values that are at a balanced level. The precision value of 0.9911 indicates that about 99.1% of all positive predictions generated by the model are correct predictions. Meanwhile, the recall value of 0.991 indicates that the model is able to identify

about 99.1% of all positive data actually contained in the dataset. The F1-score value of 0.9910, which is the harmonic mean between precision and recall, illustrates the optimal balance between the model's ability to recognize positive classes and avoid misclassification. Overall, these results show that the model has excellent generalization ability to new data and does not show any significant signs of overfitting. Figure 10 displays the Confusion Matrix illustrating the results of the model evaluation in the testing phase.

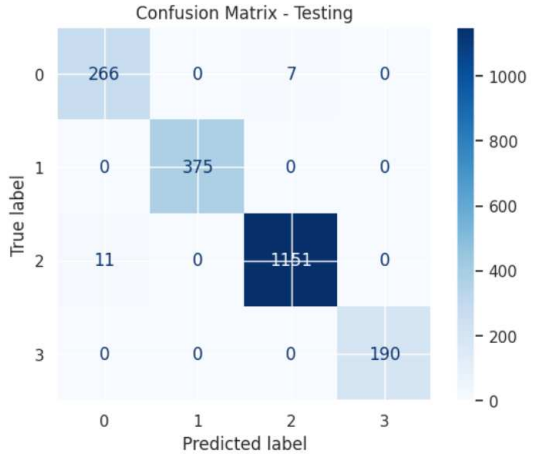

Figure 10. Confusion Matrix of Testing Data

Figure 10 displays the Confusion Matrix visualization results showing the distribution of model predictions in each sleep quality category. In the Excellent (0) class, out of a total of 273 data, 266 data were classified correctly, while 7 data were misclassified into the Good (2) class. For the Fair (1) class, all 375 data were classified correctly without any errors. For the Good (2) class, out of 1162 data, 1151 data were correctly detected, while 11 data were incorrectly identified as the Excellent (0) class. As for the Poor (3) class, all 190 data were correctly classified by the model. These results show a very high level of accuracy across all categories, with only a few misclassifications occurring in the Excellent and Good classes.

In addition, the results of the model performance evaluation were also analyzed through the ROC graph measured by the AUC value. All classes, both in the training and testing data, showed an AUC value = 1.00, indicating a very good Area Under the Curve. The ROC curve for each class was ideally positioned in the upper left corner of the graph (TPR = 1, FPR = 0), indicating that the model made almost no errors in distinguishing between positive and negative classes. The consistency of this pattern in both stages (training and testing) indicates that the model does not suffer from overfitting and is able to maintain a stable high performance. Thus, the model proved to have highly accurate classification capabilities without any significant overlap between classes. The resulting ROC graph is shown in Figure 11

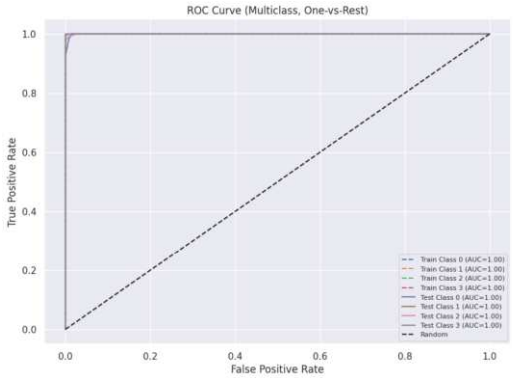

Figure 11. ROC Graph of Sleep Quality Prediction

The curve above shows that the model is able to distinguish between positive and negative classes with perfect AUC values across all classes in both training and testing data. Further analysis of these results indicates that the dataset has a clear structure and highly informative features, allowing the model to maintain high performance on the testing data. This also demonstrates the effectiveness of the preprocessing that has been applied, while providing insight into the potential generalizability of the model to new complex data.

Model Deployment

The model deployment stage aims to implement the sleep quality prediction model that has been developed into the application environment so that it can be utilized functionally. In this research, the deployment process is carried out on a local server using a web-based interface. The model that has gone through the training process is stored in a standardized format using joblib in Python, then integrated into an application that allows users to enter data directly and obtain sleep quality level prediction results automatically. The interface of the application is presented in Figure 12.

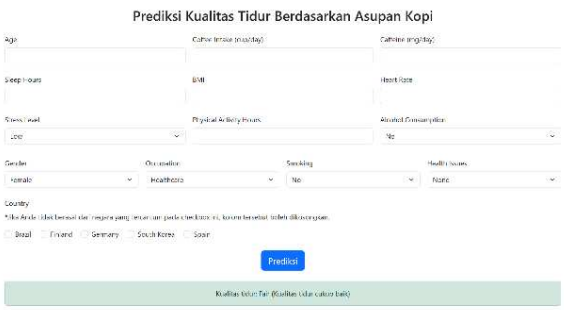

Figure 12. Application View

Internal evaluation is done descriptively to ensure the functionality and performance of the model is as expected. Details of the tasks that have been performed at this deployment stage are presented in Table 4.

Table 4. Checklist of Deployment Process and Internal Evaluation of Sleep Quality Prediction Model

No.	Task	Status
1	Model saved in a standardized format (joblib)	Completed
2	Model integrated into a web-based application	Completed
3	Application interface designed for intuitive data input and prediction output	Completed
4	Prediction process runs efficiently and consistently	Completed

CONCLUSIONS AND SUGGESTIONS

Conclusion

This research successfully developed and implemented a sleep quality prediction model with very high performance, indicated by an accuracy value of 0.991 and balanced precision, recall, and F1-score metrics across classes. The challenge of unbalanced class distribution in the initial data was overcome through the application of Synthetic Minority Oversampling Technique (SMOTE), which played an important role in improving the accuracy and stability of the model across sleep quality categories (Excellent, Fair, Good, and Poor). The Confusion Matrix evaluation results showed a very low misclassification rate, while the Area Under the Curve (AUC) value of 1.00 in all classes confirmed the model's ability to distinguish between positive and negative classes perfectly without any indication of overfitting. The learning curve analysis shows that the model has achieved optimal stability and generalization, indicating an effective learning process. In addition, the developed model has been successfully deployed in the form of a web-based sleep quality prediction application, so that it can be used practically to support sleep health analysis and monitoring. This research closes the gap related to the development of a stable and accurate model, as well as providing an applicative contribution for further testing on external datasets and comparison with other baselines.

Suggestion

This research still has the potential to be developed further. The use of larger and varied datasets can be an important step to improve the generalization ability of the model to various individual characteristics. In terms of implementation, the system can be expanded through cloud-based integration or mobile application development, so that the model is able to provide real-time predictions and reach more users. In addition, further research can also focus on expanding the features of the model, for example by adding other physiological data so that sleep quality prediction can be done more comprehensively and accurately. During development, ethical and practical aspects should be considered, including the privacy and security of health data and the interpretation of predictions to avoid errors.

REFERENCES

Agustiana P, L., & Nafisah, K. D. (2024). Hubungan Konsumsi Kopi dengan Kualitas Tidur pada Remaja. *Colostrum Mother Journal*, *1*(01), 1–23.

Firdaus, D., Sumardi, I., & Chazar, C. (2025). *Deteksi Serangan Pada Jaringan Internet Of Things Medis Menggunakan Machine Learning Dengan Algoritma XGBoost Medical of Things Attack Detection On Internet Medical Of Things Using Machine Learning With Xgboost Algorithm*. *8*(1), 34–42.

Henrich, L. C., Antypa, N., & Van den Berg, J. F. (2023). Sleep quality in students: Associations with psychological and lifestyle factors. *Current Psychology*, *42*(6), 4601–4608. https://doi.org/10.1007/s12144-021-01801-9

Khasanah, N., Eka Saputri, D. U., Aziz, F., & Hidayat, T. (2025). Studi Perbandingan Algoritma Random Forest dan K-Nearest Neighbors (KNN) dalam Klasifikasi Gangguan Tidur. *Computer Science (CO-SCIENCE)*, *5*(1), 17–25. https://doi.org/10.31294/coscience.v5i1.5522

Mansur, A. R., Farlina, M., & Paraswati, S. I. (2025). *Dunia Dalam Secangkir Kopi* (I. M. Sari (ed.); 1st ed.). Karya Bakti Makmur (KBM) Indonesia.

Mawardi, A. B., Pradini, R. S., & Haris, M. S. (2024). Komparasi Algoritma Boosting Untuk Prediksi Gangguan Tidur. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, *13*(3).

Meiranny, A., & Chabibah, A. M. (2022). Pengaruh Konsumsi Minuman Berkafein Terhadap Pola dan Kualitas Tidur Mahasiswa : A Literatur Review. *Media Publikasi Promosi Kesehatan Indonesia (MPPKI)*, *5*(2), 117–122. https://doi.org/10.56338/mppki.v5i2.1910

Mujiyono, S., Sanjaya, U. P., Wibisono, I. S., & Setyowati, H. (2025). Prediksi Fluktuasi Berat Badan Berdasarkan Pola Hidup Menggunakan Model XGBoost dan Deep Learning. *Jurnal Algoritma*, *22*(1), 221–233. https://doi.org/10.33364/algoritma/v.22-1.2253

Nawawi, I., & Fatah, Z. (2024). Penerapan Decision Trees dalam Mendeteksi Pola Tidur Sehat Berdasarkan Kebiasaan Gaya Hidup. *Jurnal Ilmiah Sains Teknologi Dan Informasi*, *2*(4), 34–41. https://doi.org/10.59024/jiti.v2i4.969

Nugraha, R., Muflih, S. R., Ferianda, I., Arashi, Z. R., Sahubawa, N. S., Hendrian, Y., & Kinanti, S. L. (2025). PENGEMBANGAN WEBSITE KLASIFIKASI KUALITAS TIDUR DAN REKOMENDASI PENANGANAN MENGGUNAKAN LOGISTIC REGRESSION. *Jurnal Riset, Inovasi Dan Teknologi Kabupaten Batang*, *9*(2), 30–36.

Permata, Y. N., Sriwiyati, K., & Affanin, S. (2023). Hubungan kebiasaan minum kopi dan aktivitas fisik dengan kualitas tidur mahasiswa fakultas kedokteran. *Journal of Nursing Practice and Education*, *4*(1), 206–211. https://doi.org/10.34305/jnpe.v4i1.957

Priambodo, A. R., & Chozanah, R. (2020). *Tergantung Metabolisme Tubuh, Efek Samping Minum Kopi Setiap Orang Beda!* Suara.Com. https://www.suara.com/health/2020/06/07/080418/tergantung-metabolisme-tubuh-efek-samping-minum-kopi-setiap-orang-beda

Putra, J. L., & Hidayat, W. F. (2024). Prediksi Kualitas Tidur: Pendekatan Machine Learning yang Mengintegrasikan Faktor Kesehatan dan Lingkungan. *Computer Science (CO-SCIENCE)*, *4*(2), 157–162. https://doi.org/10.31294/coscience.v4i2.4737

Ranti, N. B. P., Boekoesoe, L., & Ahmad, Z. F. (2022). Kebiasaan Konsumsi Kopi, Penggunaan Gadget, Stress dan Hubungannya dengan Kejadian Insomnia pada Mahasiswa. *Jambura Journal of Epidemiology*, *1*(1), 20–28. https://doi.org/10.37905/jje.v1i1.15027

Sari, A. P., Billy, B., Tsaqif, D. A., Sartono, B., & Firdawanti, A. R. (2024). Classification of Drinking Water Source Suitability in West Java Using XGBoost and Cluster Analysis Based on SHAP Values. *Indonesian Journal of Statistics and Its Applications*, *8*(2), 202–214. https://doi.org/10.29244/ijsa.v8i2p202-214

Sari, D. (2024). Prediksi Gangguan Tidur pada Sleep Health and Lifestyle Menggunakan Support Vector Machine dan Neural Network. *JAVIT : Jurnal Vokasi Informatika*, 36–42. https://doi.org/10.24036/javit.v4i1.168

Sastra, A., & Sabri, K. (2025). *Analisis Data Science terhadap Kualitas Tidur dan Faktor-Faktor yang Mempengaruhinya : Studi Kasus Sleep Health and Lifestyle Dataset*. *3*(2), 92–99.

Thanri, Y. Y., Riza, B. S., Iriani, J., & Noor, A. A. (2025). *Klasifikasi Gangguan Tidur pada Individu Menggunakan Algoritma Naive Bayes Berbasis Data Gejala Klinis*. *13*(1).

Tinaliah, T., & Elizabeth, T. (2024). Prediksi Jenis Kanker Payudara Menggunakan Metode Support Vector Machine Berbasis Recursive Feature Elimination. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, *11*(3), 1–9.

Tresnawulan, S., Afrina, R., & Kamilah, S. (2024). *Hubungan Perilaku Konsumsi Kopi dengan Kualitas Tidur pada Remaja di Kopi Janji Jiwa Depok Dua Timur Tahun 2022*. *3*.