# Prediction of Sleep Disorder using Artificial Neural Networks


1st Maesa Ken Neobi Arief
*School of Computer Science*
*Bina Nusantara University*
Jakarta, Indonesia 12410
maesa.arief@binus.ac.id

2nd Stephen Jo
*School of Computer Science*
*Bina Nusantara University*
Tangerang, Indonesia 15147
stephen.jo@binus.ac.id

3rd Diana
*Computer Science Department*
*School of Computer Science*
*Bina Nusantara University*
Jakarta, Indonesia 11480
diana@binus.edu

4th Afdhal Kurniawan.
*Computer Science Department*
*School of Computer Science*
*Bina Nusantara University*
Jakarta, Indonesia 11480
afdhal.kurniawan@binus.ac.id



***Abstract*—Sleep quality is a general issue that is desired to be achieved by most people in the world. Thus, the invention of new methods of classifying and correlating the effects of general health and wellbeing to sleep quality is needed. This paper describes the implementation of artificial neural networks in detecting the correlation of general health and sleep quality, how the performance is throughout the use of the model, and how this model will reflect the future implementation of machines in healthcare. Using certain epochs and test-size, this research is able to create an accurate detector. With the help of hyperparameter tuning, the features can become a bigger factor in the model. The output of this research is a detector with an accuracy of 88% and a MSE score of 0.0942. A sleep disorder detector using Artificial Neural Networks is successful and maintains great performance proven by the training loss and validation loss metric after each epoch.**



***Keywords*—Artificial Neural Networks, Classification, Evaluation Metrics, Health, Sleep Quality,**


## I. Introduction

Sleep quality is a desired trait by people across the world.Numerous researches have proven that there are significant and predictable patterns between sleep quality and health. Moreover, those same findings have also proven that there are not any related patterns between sleep quantity and sleep quality [1].

Poor sleep quality has a negative impact in different areas related to physical health along with other serious consequences. Other than physical health, poor sleep quality also can negatively impact mental health and psychological health such as anxiety, depression, aggression, hyperactivity disorder, and other mental health disorders [2]. Akhlagh et al. found that from a sample size of 1000 adolescents from non-private schools in Iran, 56% were grouped as poor sleepers. And according to the results in the study, 42% of students had poor general health [3].

To determine a solution for this problem, it would be necessary to detect a correlation using automated methods. However, implementing these methods requires comprehensive knowledge about the effects of the lack of sleep quality in accordance with general health and well-being due to sleep disorders [4]. In order to achieve this, an implementation of a machine-learning based system to predict the outcome of general health and wellbeing using sleep quality is required to assist in identifying common trends and patterns in sleep. A prediction system can collect data and use it to determine a correlation [5].

The predictor will implement the ANN (Artificial Neural Networks) model. Artificial Neural Networks imitates the principles of the human brain and computes quantitative data and creates a mathematical model of biological neurons. The reason why this model is chosen is due to the complex amount of quantitative data in the dataset and the multivariate algorithm that the ANN model provides [6].

The superiority of ANN over other methods for health prediction lies in its ability to capture complex non-linear relationships within data, particularly in the healthcare domain. Artificial Neural Networks, such as the proposed artificial backpropagation scaled conjugate gradient neural network algorithm, provide significant advantages in terms of predictive accuracy and efficiency. In addition, ANN can handle the intricate physiological characteristics inherent in medical data, making them well-suited for tasks like health prediction. [15]

## II. Related Works

Sleep quality is a crucial aspect that has a strong correlation with physical and psychological well-being, as well as with the risk of death. People who have poor sleep quality are more likely to experience distress, depression, anxiety, and general health problems [7]. Over the past few decades, extensive research has revealed a complex relationship between sleep quality and various health dimensions in children and adolescents. Short sleep duration has been linked to daytime sleepiness, obesity, specific sleep disorders, and various health characteristics. In addition, sleep duration has been associated with high blood pressure, pain perception, racial disparities, and cognitive function. Studies have also found that subjective psychological well-being, lifestyle habits, and certain risk-taking behaviors, such as binge drinking, are related to sleep quality and may contribute to negative health outcomes [8].

Artificial Neural Networks (ANN) can be attributed to their potential to revolutionize health prediction, especially in assessing sleep quality and its impact on overall well-being. By modeling complex relationships and analyzing patterns within sleep data, ANN can present a hopeful path for forecasting health consequences linked to different levels of sleep quality [9].

In a study investigating the use of machine learning (ML) to detect sleep apnea (SA), researchers analyzed data from 620 patients referred to a suburban community sleep center. The dataset included variables such as height, weight, waist, hip, BMI, age, neck size, Modified Friedman (MF), snoring, Epworth sleepiness scale, sex, and daytime sleepiness. The proposed method utilized a binary particle swarm optimization (BPSO) technique to select the best features characterizing apnea, followed by developing an Artificial Neural Network (ANN) model based on the feedforward algorithm. The data was clustered by gender, resulting in four different models with and without feature selection. A total of 93 experiments indicated that neck size and daytime sleepiness were the most valuable features, selected 86% and 87% of the time, respectively, while MF was weakly associated and chosen 44% of the time. The ANN model achieved a detection ratio of 80% for men and 75% for women. The study concluded that ML algorithms, with their advanced feature selection capabilities, outperform traditional linear approaches in solving complex problems like sleep apnea, highlighting the need for further investigation into ML-based methods for SA detection [22].

## III. MATERIAL AND METHOD

### A. Dataset

Quality of the dataset serves as the cornerstone of machine-learning research, facilitating training and testing of models while providing stable points of comparison through benchmark datasets. They not only coordinate researchers around shared problems but also represent meaningful abstractions of real-world tasks of problem domains. Thus, ensuring alignment with real-world tasks is crucial for the accurate measurement of scientific progress and the safe deployment of the models [17].

The dataset that this paper uses is from Kaggle and contains 374 rows of data, with each row containing 13 columns: Person ID, Gender, Age, Occupation, Sleep Duration, Quality of Sleep, Physical Activity Level, Stress Level, BMI Category, Blood Pressure, Heart Rate, Daily Steps, and Sleep Disorder.

### B. Data Preprocessing

In the realm of real-world data, encountering instances of dirtiness, incompleteness, and inconsistency is commonplace. The role of data preprocessing techniques becomes imperative, aiming to enhance data quality and subsequently bolster the accuracy and efficiency of subsequent mining processes. Data preprocessing holds significance as it predicates quality decisions upon quality data. The early detection and rectification of data anomalies, alongside the reduction of the dataset to be analyzed, yield substantial dividends in facilitating informed decision-making processes [16].

To begin the data preprocessing phase, we cleaned the data by removing any missing or irrelevant values from the columns in our dataset. Since the dataset consists of 374 rows of data and with 13 columns for each row, this step was crucial to ensure the accuracy of our analysis.

Next, we transformed the categorical data(such as Gender, Occupation, and Sleep Duration) into numerical data, which makes it easier to analyze using machine learning techniques. After that, we normalized the numerical data, such as Sleep Duration, Physical Activity Level, Stress Level, BMI Category, Blood Pressure, Heart Rate, and Daily Steps, to ensure all the data was on a consistent scale.

In addition, the categorical variables are label encoded such as Gender, BMI Category, and Sleep Disorder. The utilization of label encoding is warranted due to its effectiveness in transforming categorical data into numerical representations, thus enabling machine learning algorithms to process and interpret such data more efficiently. Moreover, we plotted the data to visually explore within the dataset, providing a deeper analysis to uncover underlying patterns and insights.

Finally, we split the preprocessed data into training and testing data, ensuring that the data was balanced and each data contained a representative sample of the original data. This allowed us to train and validate our model working effectively and assess their performance on unseen data.

Overall, these data preprocessing steps were crucial in preparing the data for analysis and training our model to be accurate, reliable, and reproducible.

### C. Methods

Hyperparameter tuning involves the selection of optimal values for parameters that are not learned during model training. These parameters, such as the learning rate, number of layers, and number of neurons, have a significant impact on the model's performance [18].

Evaluation metrics, on the other hand, are employed to assess the model's performance by quantifying its predictive accuracy. These metrics are crucial in achieving an optimal classifier during classification training. Common evaluation metrics include accuracy, precision, recall, F1-score, Mean Square Error (MSE), and area under the ROC curve (AUC) [19].

#### a. Mean Squared Error (MSE)

Mean Square Error (MSE) calculates and measures the difference between predicted solutions and desired solutions. The smaller the MSE value is, the better performance the classification model has. The formula for MSE is defined below:

$$MSE = \frac{1}{n}\sum_{j=1}^{n}(Pj - Aj)^2 \tag{1}$$

where Pj is a predicted value of instance j, Aj is the desired value of instance j, and n is the number of instances. The more minimum score of MSE, the better the model's performance is [19].

#### b. Area Under the ROC Curve (AUC)

Area under the ROC Curve (AUC) is widely used to construct an optimized learning model and to compare learning algorithms. The AUC value visualizes the overall ranking performance for a classifier model. The formula for AUC is defined below:

$$AUC = \frac{S_p - n_p(n_n + 1)/2}{n_p n_n} \tag{2}$$

#### c. Confusion Matrix

A confusion matrix is generally used to measure the performance of a classification algorithm (supervised model) by comparing the predicted value to the true value. The idea is that it forms its own equivalent classes of the data, and assigns the classes to the categories provided by the dataset. By using the confusion matrix, the other evaluation metrics can be calculated using the data gathered from the matrix, such as accuracy, recall, and precision [21].

| | | True value | |
|---|---|---|---|
| | | $P$ | $N$ |
| Predicted value | $\hat{P}$ | True Positive | False Positive |
| | $\hat{N}$ | False Negative | True Negative |

Fig. 1. A 2-class confusion matrix

### d. Artificial Neural Networks

Artificial Neural Networks (ANNs) are artificial intelligence models that are derived from neurons in the human body. ANNs consist of interconnected nodes, or neurons, organized in layers, and they are widely used in various machine learning tasks. Artificial Neural Networks are highly capable of recognizing patterns and making decisions to create robust classifiers to solve non-linear problems [20].

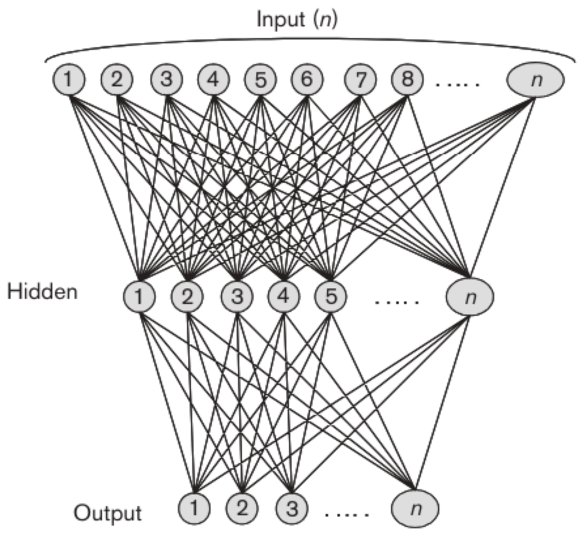


Fig.2. Structure of an ANN

Neurons are often organized in a topological manner, depending on the amount of input data. Then, information is moved from one hidden layer to another hidden layer in the ANN. The output layer acts as the last layer in the structure which provides the final result. The output layer contains only one processing element. All processing elements are connected to a weight which modifies the input and output value [20].

## IV. RESULT

The performance of the model is evaluated using specific metrics. The testing was conducted with a test size of 30% and a random state set to 42 to ensure reproducibility. The Sequential Artificial Neural Network (ANN) model comprised an input layer with 12 neurons, a hidden layer with 8 neurons and an output layer with 1 neuron, all activated using Rectified Linear Unit (ReLU) and sigmoid functions respectively.

The model was compiled with the binary cross-entropy loss function and optimized using the Adam optimizer. To avoid overfitting, L1/L2 regularization was applied to the weights of the neural network layers with regularization coefficients set to 0.03 for both L1 and L2. The performance of the model was evaluated using several metrics, including accuracy, precision, recall and Mean Squared Error (MSE).

Training of the model was executed with 75 epochs and a batch size of 15. Upon evaluation on the dataset, the model achieved an accuracy roughly 88%. Additionally, the Mean Squared Error was monitored throughout the training process to assess the model's performance. The MSE gradually decreased over the epochs, reaching a final value of 0.0942, indicating that the model's predictions closely matched the actual values in the test data.

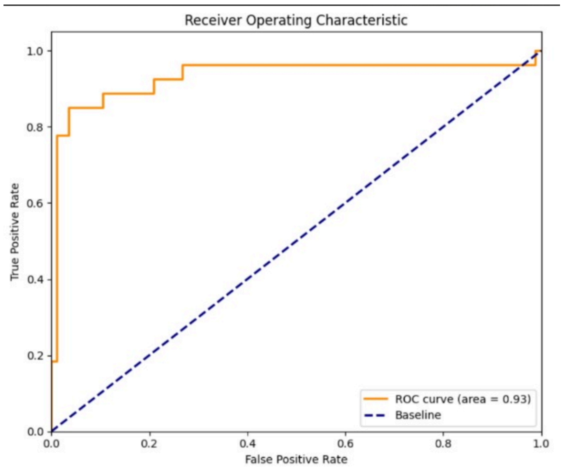


Fig 3. ROC Curve

Fig 3. Shows the Receiver Operating Characteristic (ROC) curve, while illustrates the model's ability to distinguish between positive and negative classes. The area under the ROC curve (AUC) is 0.93, indicating a high level of distinction. An AUC value close to 1 signifies that the model is excellent at distinguishing positive and negative classes.

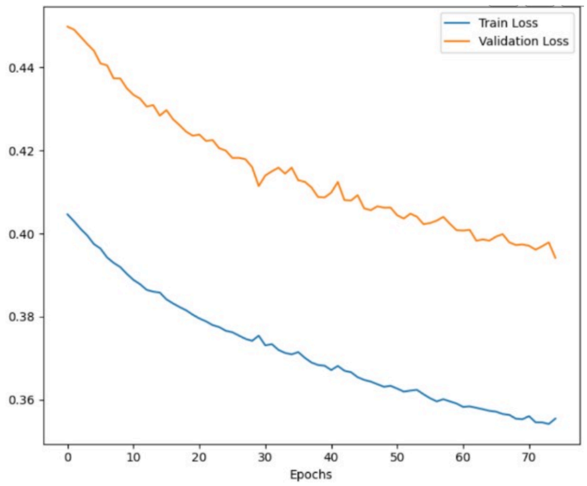


Fig 4. Loss Over Epochs

Fig 4. Shows the loss during training and validation over 75 epochs. The training loss, represented by the blue line, shows a consistent decrease, indicating that the model is learning and improving its fit on the training data. The validation loss, represented by the orange line, also shows a decrease, little with fluctuations, indicating that the model generalizes well to unseen data without significant overfitting. These results demonstrate that the model effectively learns the patterns of the data and maintains good performance on the validation data.

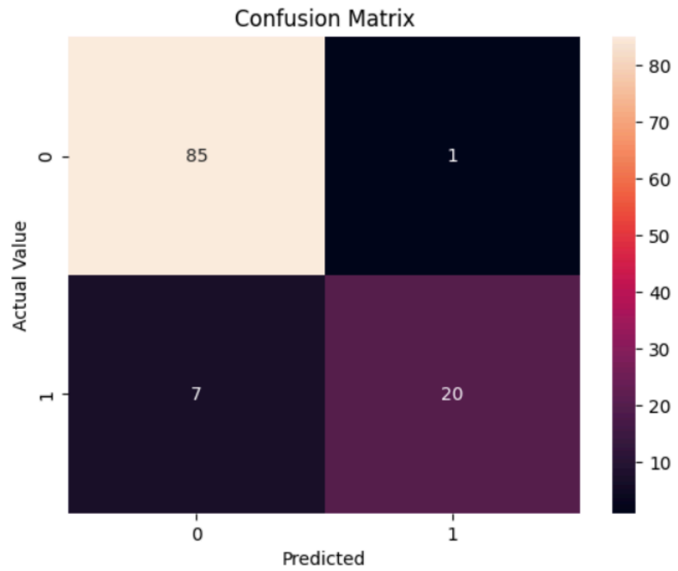


Fig 5. Confusion Matrix

Fig 5. Provides a summary of prediction results on classification tasks. It shows the counts of True Positive, True Negative, False Positive, and False Negative. The results indicate that the model correctly predicted 20 cases (True Positive) where the positive class was present and correctly identified 85 cases (True Negative) where the negative class was present. However, there were 7 cases (False Positive) where the model incorrectly identified the positive class, and 1 case (False Negative) where the model missed predicting the positive class.

## V. CONCLUSION

Further research in accordance with this topic will be needed to improve the model and make this a future topic in the biomedical industry and give a new urgency to health and sleep quality. The research done based on this model is aimed to find the most contributing factors through machine learning feature vectors and see what really factors into the quality of sleep.

Artificial Neural Networks (ANN) is an effective and reasonable approach to our research because of the capability of the model to handle multivariate variables, however it also has some limitations. In most cases, classification with more than two variables limits the model's performance. During research, the use of three labels affects the model's performance. Other than that, if there are large amounts of data, it takes time to process the model and can lead to large amounts of memory taken from the compiler. And the ANN model is prone to overfitting because it requires specific datasets and testing sizes to be effective.

## REFERENCES


[1] J. J. Pilcher, D. R. Ginter, and B. Sadowsky, "Sleep Quality versus Sleep quantity: Relationships between Sleep and Measures of health, well-being and Sleepiness in College Students," Journal of Psychosomatic Research, vol. 42, no. 6, pp. 583–596, Jun. 1997, doi: https://doi.org/10.1016/s0022-3999(97)00004-4.

[2] V. Clement-Carbonell, I. Portilla-Tamarit, M. Rubio-Aparicio, and J. J. Madrid-Valero, "Sleep Quality, Mental and Physical Health: a Differential Relationship," International Journal of Environmental Research and Public Health, vol. 18, no. 2, p. 460, Jan. 2021, doi: https://doi.org/10.3390/ijerph18020460.

[3] A.-A. K. Akhlagh and M. F. Ghalebandi, "Sleep Quality and Its Correlation with General Health in Pre-university Students of Karaj, Iran," Iranian Journal of Psychiatry and Behavioral Sciences, vol. 3, no. 1, pp. 44–9, 2009, Accessed: Apr. 10, 2024. [Online]. Available: https://brieflands.com/articles/ijpbs-2842

[4] S. Chokroverty, "Overview of Sleep & Sleep Disorders," Indian Journal of Medical Research, vol. 131, no. 2, p. 126, Feb. 2010, Available: https://journals.lww.com/ijmr/Abstract/2010/31020/Overview_of_sleep___sleep_disorders.4.aspx

[5] V. Farrahi and M. Rostami, "Machine Learning in Physical activity, sedentary, and Sleep Behavior Research," Journal of Activity, Sedentary and Sleep Behaviors , vol. 3, no. 1, Jan. 2024, doi: https://doi.org/10.1186/s44167-024-00045-9.

[6] A. Duykuluoglu, "The Significance of Artificial Neural Networks in Educational Research a Summary of Research and Literature | Technium BioChemMed," techniumscience.com, vol. 2, no. 2, Jun. 2021, Accessed: Jun. 02, 2024. [Online]. Available: https://techniumscience.com/index.php/biochemmed/article/view/3818

[7] A.-A. K. Akhlagh and M. F. Ghalebandi, "Sleep Quality and Its Correlation with General Health in Pre-university Students of Karaj, Iran," Iranian Journal of Psychiatry and Behavioral Sciences, vol. 3, no. 1, pp. 44–9, 2009, Available: https://brieflands.com/articles/ijpbs-2842

[8] T. Paiva, T. Gaspar, and M. G. Matos, "Sleep deprivation in adolescents: correlations with health complaints and health-related quality of life," Sleep Medicine, vol. 16, no. 4, pp. 521–527, Apr. 2015, doi: https://doi.org/10.1016/j.sleep.2014.10.010.

[9] N. Malik, "Artificial Neural Networks and their Applications," arXiv.org, May 10, 2005. https://arxiv.org/abs/cs/0505019 (accessed Jun. 25, 2023).

[10] S. Cokim and J. Ghaly, "Using Machine-Learning to Predict Sleep-Disordered Breathing Diagnosis from Medical Comorbidities and Craniofacial Features," Cureus, vol. 15, no. 5, May 2023, doi: https://doi.org/10.7759/cureus.39798.

[11] M. Santosh A. Shinde and D. P. Raja Rajeswari, "Intelligent Health Risk Prediction Systems Using Machine learning: a Review," International Journal of Engineering & Technology, vol. 7, no. 3, p. 1019, Jun. 2018, doi: https://doi.org/10.14419/ijet.v7i3.12654.

[12] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and Prospects," Science, vol. 349, no. 6245, pp. 255–260, Jul. 2020, doi: https://doi.org/10.1126/science.aaa8415.

[13] M. Gevrey, I. Dimopoulos, and S. Lek, "Review and Comparison of Methods to Study the Contribution of Variables in Artificial Neural Network Models," Ecological Modelling, vol. 160, no. 3, pp. 249–264, Feb. 2003, doi: https://doi.org/10.1016/s0304-3800(02)00257-0.

[14] D. Gopu, "Prediction of Diseases in Smart Health Care System Using Machine Learning," International Journal of Recent Technology and Engineering, vol. 8, no. 5, pp. 2534–2537, Jan. 2020, doi: https://doi.org/10.35940/ijrte.e6482.018520.

[15] M. M. Bukhari, B. F. Alkhamees, S. Hussain, A. Gumaei, A. Assiri, and S. S. Ullah, "An Improved Artificial Neural Network Model for Effective Diabetes Prediction," Complexity, vol. 2021, pp. 1–10, Apr. 2021, doi: https://doi.org/10.1155/2021/5525271.

[16] T. Iliou, C. Anagnostopoulos, M. Nerantzaki, and G. Anastassopoulos, "A Novel Machine Learning Data Preprocessing Method for Enhancing Classification Algorithms Performance," in International Conference on Engineering Applications of Neural Networks, Sep. 2015.

[17] B. Koch, E. Denton, A. Hanna, and J. G. Foster, "Reduced, Reused and Recycled: the Life of a Dataset in Machine Learning Research," in 35th Conference on Neural Information Processing Systems , Cornell University, Jan. 2021. doi: https://doi.org/10.48550/arxiv.2112.01716.

[18] T. Yu and H. Zhu, "Hyper-Parameter Optimization: A Review of Algorithms and Applications," arXiv.org, Mar. 12, 2020. https://arxiv.org/abs/2003.05689

[19] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," International Journal of Data Mining & Knowledge Management Process, vol. 5, no. 2, pp. 01-11, Mar. 2015, doi: https://doi.org/10.5121/ijdkp.2015.5201.

[20] E. Grossi and M. Buscema, "Introduction to Artificial Neural Networks," European Journal of Gastroenterology & Hepatology, vol. 19, no. 12, pp. 1046–1054, Dec. 2007, doi: https://doi.org/10.1097/meg.0b013e3282f198a0.

[21] I. Düntsch and G. Gediga, "Confusion Matrices and Rough Set Data Analysis," Journal of Physics: Conference Series, vol. 1229, p. 012055, May 2019, doi: https://doi.org/10.1088/1742-6596/1229/1/012055.

[22] S. Surani, A. Sheta, H. Turabieh, J. Park, S. Mathur, and A. Katangur, "Diagnosis of Sleep Apnea Using Artificial Neural Network and Binary Particle Swarm Optimization for Feature Selection," Chest, vol. 156, no. 4, p. A136, Oct. 2019, doi: https://doi.org/10.1016/j.chest.2019.08.215.


## ACKNOWLEDGEMENTS


This work has been created for the purpose of fulfilling the final project of the subject: Research Methodology in Computer Science, as well as an attempt to publicize our work to the research world. We also would like to thank Mrs. Diana and Mr. Afdhal Kurniawan as our lecturers for their support and helpful comments.