

MACHINE LEARNING PREDICTION OF SLEEP QUALITY AMONG UITM PERLIS STUDENTS USING THE PITTSBURGH SLEEP QUALITY INDEX

Nur Aisah Binti Mohd Rosdi, Nor Hayati Binti Shafii

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Perlis Branch, Malaysia

e-mail: nurraisah09@gmail.com, norhayatishafii@uitm.edu.my

Abstract

Sleep serves as a fundamental process for repairing the body, replenishing mental energy, and maintaining emotional stability (Zhang et al., 2023). However, a low quality of sleep is typical of the university students. This study aimed to assess sleep quality among UiTM Perlis students using machine learning (ML) and to compare the performance of XGBoost, Random Forest, and Decision Tree models in predicting sleep quality. A total of 242 responses were collected using the Pittsburgh Sleep Quality Index (PSQI) questionnaire (Buysse et al., 1989), which included demographic and lifestyle factors. After encoding, scaling, and preprocessing, all models were tuned using GridSearchCV with 5-fold cross-validation. Performance was evaluated using accuracy, precision, recall, F1-score, and AUC. The mean PSQI global score was 7.18, indicating that most students had poor sleep quality. Among the models tested, XGBoost achieved the highest test accuracy (97.96%) and best cross-validation score (95.34%) with the lowest standard deviation, confirming its accuracy and stability. Overall, this study demonstrates that ML models, particularly XGBoost, can effectively predict sleep quality among UiTM Perlis students.

Keywords: *Sleep Quality, Pittsburgh Sleep Quality (PSQI), Machine Learning, Prediction*

1. Introduction

Adequate sleep is universally recognised as a key factor in sustaining optimal health and performance (Zhang et al., 2023). According to Hirshkowitz et al. (2015), adults aged 18 to 64 are recommended to get 7 to 9 hours of sleep each night. However, poor sleep quality is prevalent among university students due to academic pressure, irregular routines, and excessive screen time (Bitkina et al., 2022). These issues may lead to fatigue, stress, and reduced academic performance. This study addresses by applying machine learning (ML) techniques to predict sleep quality among UiTM Perlis students. The main objectives are to predict students' sleep quality using ML models and to compare the performance of XGBoost, Random Forest, and Decision Tree algorithms based on PSQI data collected from UiTM Perlis.

2. Methodology

This study applied machine learning to predict sleep quality among UiTM Perlis students using PSQI data and selected demographics. A total of 242 valid responses were collected via an online questionnaire covering seven PSQI components and demographic details. Preprocessing steps included handling missing values, one-hot encoding, and standard scaling. Models were developed using Python (scikit-learn, XGBoost) and optimised with GridSearchCV using 5-fold cross-validation. Performance was evaluated using accuracy, precision, recall, F1 score, and AUC, ensuring a systematic and reproducible prediction pipeline.

3. Results / Findings

The outcomes of this study demonstrated that machine learning techniques are effective in predicting sleep quality among UiTM Perlis students. Among the three models evaluated, Decision Tree,

Random Forest, and XGBoost. The XGBoost model showed the best overall performance, achieving the highest accuracy (97.96%) and AUC. This indicates that XGBoost was most effective in capturing complex patterns within the PSQI and demographic data. The models were evaluated using key classification metrics including accuracy, precision, recall, F1-score, and AUC. Overall, these results highlight that machine learning approaches, particularly XGBoost, are well-suited for accurate and efficient assessment of sleep quality.

4. Novelty of Research / Product / Project

This study introduces a novel approach by applying machine learning (ML) techniques to predict sleep quality among university students using the PSQI and selected demographic variables. While previous studies relied solely on descriptive or inferential analysis of PSQI, (Farah et al., 2019). This research incorporates supervised ML models to develop predictive insights, enabling early identification of poor sleep quality. The integration of PSQI components with ML enhances the objectivity and reproducibility of sleep quality assessment. This work contributes to the growing field of health informatics by demonstrating how machine learning can support data-driven interventions to improve student well-being.

5. Conclusion or Future Work

This study demonstrated the effectiveness of machine learning, particularly XGBoost, in predicting sleep quality using PSQI data. Future research could explore larger and more diverse datasets, include objective sleep measures (e.g., wearable devices), and apply deep learning models to enhance prediction accuracy and generalisability.

REFERENCE LIST

- Bitkina, O. V., Park, J., & Kim, J. (2022). Modeling sleep quality depending on objective actigraphic indicators based on machine learning methods. *International Journal of Environmental Research and Public Health*, 19(16), 9890. <https://doi.org/10.3390/ijerph19169890>
- Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2), 193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)
- Farah, N. M., Yee, T. S., & Rasdi, H. F. M. (2019). Self-Reported Sleep quality using the Malay version of the Pittsburgh Sleep Quality Index (PSQI-M) in Malaysian adults. *International Journal of Environmental Research and Public Health*, 16(23), 4750. <https://doi.org/10.3390/ijerph16234750>
- Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., Hazen, N., Herman, J., Katz, E. S., Kheirandish-Gozal, L., Neubauer, D. N., O'Donnell, A. E., Ohayon, M., Peever, J., Rawding, R., Sachdeva, R. C., Setters, B., Vitiello, M. V., Ware, J. C., & Hillard, P. J. A. (2015). National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health*, 1(1), 40–43. <https://doi.org/10.1016/j.sleh.2014.12.010>
- Zhang, S., Liu, X., Chen, J., Yang, H., Chen, J., Li, D., Xu, H., Wang, S., Guo, H., Zhang, N., Liu, Z., Min, X., & Wu, W. (2023). Patterns of sleep quality and its influence factors: A latent class model among students of medical university in Hubei Province, China. *Journal of Affective Disorders*, 347, 320–326. <https://doi.org/10.1016/j.jad.2023.11.090>