UNIVERSITI TEKNOLOGI MARA

I-URIID 2025

INTERNATIONAL UNDERGRADUATE RESEARCH INNOVATION, INVENTION AND DESIGN

UNDERGRADUATES PIONEERING TOMORROW'S BREAKTHROUGH

E-BOOK OF EXTENDED ABSTRACTS

SLEEP DISORDER DETECTION SYSTEM USING SUPPORT VECTOR MACHINE

Wan Seri Irisya Wan Ahmad Farhan[1], Normadiah Mahiddin[2*], Noraini Mohamed[3], Norhabibah Tarmuji[3], Rozita Mohamed[3], and Ana Salwa Shafie[3]

[1] Fakulti Sains dan Matematik, Universiti Teknologi MARA Cawangan Terengganu Kampus Kuala Terengganu
[2] Fakulti Sains dan Matematik, Universiti Teknologi MARA Cawangan Pahang Kampus Raub
[3] Fakulti Sains dan Matematik, Universiti Teknologi MARA Cawangan Pahang Kampus Jengka

*normadiah@uitm.edu.my

ABSTRACT

Sleep disorders are conditions that disrupt normal sleep patterns and can negatively affect physical, emotional, and mental well-being. These issues are becoming more widespread among adults, yet many individuals avoid diagnosis due to the high cost, discomfort, and time-consuming nature of clinical methods like polysomnography. Additionally, most machine learning systems rely on physiological data that are hard to obtain. This project focuses on developing a sleep disorder detection system using the Support Vector Machine (SVM) algorithm. The Sleep Health and Lifestyle Dataset was used, and the data underwent cleaning, encoding, normalisation, and feature selection through Recursive Feature Elimination (RFE). The dataset was then split into 60:40, 70:30, and 80:20 ratios to train and test the SVM model, which was implemented using both default and tuned versions with GridSearchCV. Performance was evaluated using accuracy, precision, recall, and F1-score, where the best results were achieved with the 60:40 split. The tuned SVM model reached an accuracy of 95.33%, precision of 95.35%, recall of 95.33%, and F1-score of 95.34%, showing strong and balanced classification between “None” and “Sleep Disorder” classes. In summary, the system acts as a useful tool for early screening using simple health inputs, although it is not a replacement for professional diagnosis and may require integration with external devices for more accurate results.

Keywords: Sleep Disorder, Diagnosis, Classification, Machine Learning, Support Vector Machine (SVM)

INTRODUCTION

Sleep is a physiological process in which the body and mind are resting by limiting some activity and interaction with the environment. The temporary withdrawal of sensory and voluntary muscles gives the body the time and opportunity to rest and gain energy. Although the processes which take place in the brain and body during sleep have not been fully explored, sleep is widely accepted as a vital organ for health (Dritsas et al., 2024). According to Alazaidah et al. (2023), sleep influences many aspects of daily life such as the way thinking, learning, working and reacting. A sleep disorder, a condition in which there is a lack of quality sleep, can disrupt these functions and activate hunger-regulating hormones, leading to increased consumption of sugary, salty, and fat-rich foods, thereby raising the risk of heart disease. Moreover, sleep disorders can lead to serious health problems such as cardiovascular diseases, obesity, and mental health issues like depression and anxiety (Dritsas et al., 2024). Thus, it is important to have a way to tell when sleep goes off track to prevent many health issues from happening later.

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that allows computer systems to learn and optimise the information without programming. It is borrowing concepts of computer science, statistics and others. The ML approach is one of supervised learning, in which the training is done with labelled data for making any prediction or decision. Classification, as one type of supervised learning, identifies boundaries that separate classes into distinct groups. Classification algorithms can make better decisions if they are allowed to lay out what decisions best fit certain boundaries (Cervantes et al., 2020). The technique that holds the opportunity to classify the presence of sleep disorder is Support Vector Machine (SVM), as it enables it to detect complex data like the duration of sleep and level of stress that are indicators to the quality of sleep. The algorithm is popularly implemented in classification because it analyses the case of the best line also referred to as the decision boundary. It begins with evaluating the training dataset, before testing it on a separate dataset to determine how well a classification algorithm can predict the outcomes. Hence, this study aims to use SVMs to develop a classification system for those who struggle with sleep disorder.

RESULTS AND DISCUSSION

This section detailed the development, evaluation, and implications of a sleep disorder classification system built using a Support Vector Machine (SVM) algorithm. The core objective was to create a robust and accurate binary classifier to distinguish individuals with a sleep disorder from those without. The system was developed in Python, utilising a structured pipeline from data preprocessing to classification. A critical design choice was the adoption of a binary classification framework, grouping all sleep disorders into a single category. This approach, endorsed by a medical expert, simplifies the model's task, mitigates class imbalance, and positions the system as a general screening tool rather than a specific diagnostic instrument, appropriately directing users to seek professional medical diagnosis for any detected condition.

Table 1: Performance Metrics Result

Split	Model Type	Accuracy (%)	Precision (None)	Precision (SD)	Recall (None)	Recall (SD)	F1-Score (None)	F1-Score (SD)
80/20	DefaultSVM	89.33	95.0	82.86	86.36	93.55	90.48	87.88
	GridSearchCV	92.00	95.24	87.88	90.91	93.02	93.02	90.62
70/30	DefaultSVM	92.04	95.24	88.00	90.91	93.62	93.62	90.72
	GridSearchCV	93.81	95.38	91.67	93.94	93.62	94.06	92.63
60/40	DefaultSVM	93.33	96.43	89.39	92.05	95.16	94.19	92.19
	GridSearchCV	95.33	96.55	93.65	95.45	95.16	96.00	94.40

To identify the optimal model configuration, rigorous testing was conducted using three different train-test data splits (60:40, 70:30, and 80:20) as shown in Table 1. For each split, two models were compared: a baseline SVM with default linear kernel settings and a tuned SVM optimised using GridSearchCV. This tuning process systematically explored combinations of kernels, regularization parameters (C), and gamma values to maximise performance. The results demonstrated conclusively that the GridSearchCV-tuned models consistently outperformed their default counterparts across all split ratios. The best results were achieved using a 60:40 data split with the tuned SVM model, which yielded an exceptional accuracy of 95.33%. The model also achieved high balanced scores in precision (95.35%), recall (95.33%), and F1-score (95.34%), indicating its proficiency in correctly identifying both classes—those with and without a sleep disorder.

This performance was benchmarked against a previous study by Alshammari et al. (2024), which used the same dataset but a more complex three-class classification and a differently tuned SVM. This project's result of 95.33% accuracy surpassed the benchmark's 92.04%, an improvement attributed to several factors: the effective binary classification strategy, the thorough hyperparameter optimisation via GridSearchCV, and a more comprehensive data preprocessing regimen that included advanced cleaning, feature encoding, scaling, and Recursive Feature Elimination (RFE). Furthermore, the selection of input features was validated through a correlation matrix, which identified strong relationships between the target variable and factors like stress level, sleep quality, and BMI. This technical validation was aligned with a list of medically relevant features suggested by a domain expert, ensuring the model's foundation was both statistically sound and clinically pertinent.

In conclusion, the research successfully fulfilled its objectives by designing, implementing, and rigorously evaluating an SVM-based classification system. The final optimized model demonstrates high accuracy and reliability, validated through comparative analysis and expert insight. The resulting prototype serves as a potent proof-of-concept for an accessible public health tool, capable of enabling early screening for sleep disorders and encouraging individuals to seek timely professional medical advice.

CONCLUSION

The sleep disorder classification system built using SVM successfully met its goals. It identifies sleep disorders based on a person’s lifestyle and health inputs. With proper data preparation and model tuning, the system reached an accuracy of 95.33%. This shows that SVM can be a helpful tool for early screening, offering a simple and straightforward solution. While there are still areas to improve, the system gives a strong starting point for future work and real-world use.

REFERENCES

Alazaidah, R., Samara, G., Aljaidi, M., Haj Qasem, M., Alsarhan, A., & Alshammari, M. (2023). Potential of machine learning for predicting sleep disorders: A comprehensive analysis of regression and classification models. *Diagnostics, 14*(1), 27. https://doi.org/10.3390/diagnostics14010027

Alshammari, T. S. (2024). Applying machine learning algorithms for the classification of sleep disorders. *IEEE Access, 12*, 36110–36121. https://doi.org/10.1109/access.2024.3374408

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing, 408*, 189–215. https://doi.org/10.1016/j.neucom.2019.10.118

Dritsas, E., & Trigka, M. (2024). Utilizing multi-class classification methods for automated sleep disorder prediction. *Information, 15*(8), 426. https://doi.org/10.3390/info15080426