AI-Based Sleep Quality Prediction – Predict sleep patterns based on fitness data.

Siddhant Udavant

***Abstract*—This project investigates the application of artificial intelligence (AI) to predict an individual's sleep quality using data collected from wearable fitness devices. By analysing variables such as heart rate, activity levels, total sleep time, environmental factors, and lifestyle habits—including caffeine and alcohol consumption—a neural network model was developed to forecast sleep quality. The model achieved a Mean Squared Error (MSE) of 0.82, indicating high predictive accuracy. A dataset comprising 100 participants over 100 days was used. This approach aims to provide personalized sleep hygiene recommendations. Results demonstrate the effectiveness of machine learning techniques in interpreting biometric and behavioral data to generate actionable health insights.**

Keywords*—Artificial Intelligence, sleep quality, wearable devices, machine learning, fitness data, neural networks, personalized recommendations, sleep hygiene, health monitoring, sleep prediction.

I. Introduction

Sleep quality is essential for overall health and wellbeing. Poor sleep is associated with various health issues including cardiovascular diseases, obesity, and mental health disorders. The advent of wearable technology has enabled the collection of detailed fitness data, presenting opportunities to monitor and analyse sleep patterns. This project explored the use of AI to predict sleep quality based on fitness data, aiming to provide personalized recommendations to improve sleep habits.

II. Research Question

Can machine-learning algorithms accurately predict sleep quality using fitness data collected from wearable devices?

III. Significance of Study

This study bridges the gap between wearable fitness technology and AI-driven health. By offering personalized sleep recommendations based on predictive models, individuals can make informed decisions to enhance sleep quality and overall health.

IV. Literature Review

Previous studies have examined the relationships between various factors and sleep quality. For example, heart rate variability has been linked to different sleep stages, whereas activity levels and screen time have been associated with sleep duration and quality. However, research on integrating multiple factors using AI to predict sleep quality remains limited.

V. Methodology

5.1 Research Design

This study adopts a quantitative research design, utilizing machine learning algorithms for predictive analysis.

5.2 Sample and Sampling Technique

The sample comprised fitness data from 100 participants over 100 days. The data included the heart rate, activity level, sleep duration, room temperature, caffeine and alcohol consumption, and screen time.

5.3 Data Collection Methods

The data were collected using wearable fitness devices and stored in a CSV file. A sample CSV file was created using Python code as follows:

import NumPy as np import pandas as pd import so defcreate_samplecsv(file_path):ifnotos.path.exists (file_path):sample_data={'timestamp':pd.date_range(start='2023-01-01', periods=100,freq='D'),'heart_rate':np.random.randint(60,100,size=100),'activity_level':np.random.randint(1000, 10000, size=100), 'sleep_duration': np.random.uniform(5,9, size=100), 'sleep_quality': np.random.uniform(1,10,size=100),'room_temperature': np.random.uniform(18, 25, size=100), 'caffeine_consumption': np.random.randint(0, 3, size=100), 'alcohol_consumption': np.random.randint(0, 3, size=100), 'screen_time': np.random.uniform(0, 3, size=100) } df = pd.DataFrame(sample_data) df.to_csv(file_path, index=False) print(f"Sample CSV file created at {file_path}") create_sample_csv('fitness_data.csv')

5.4 Data Analysis Techniques

The collected data were pre-processed to handle missing values and extract the relevant features. The features were scaled using standardization, and a neural network model was built using TensorFlow's Keros API. The model was trained, evaluated, and used to make the predictions. The data analysis process was performed using the following Python code:

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

5.4 Images

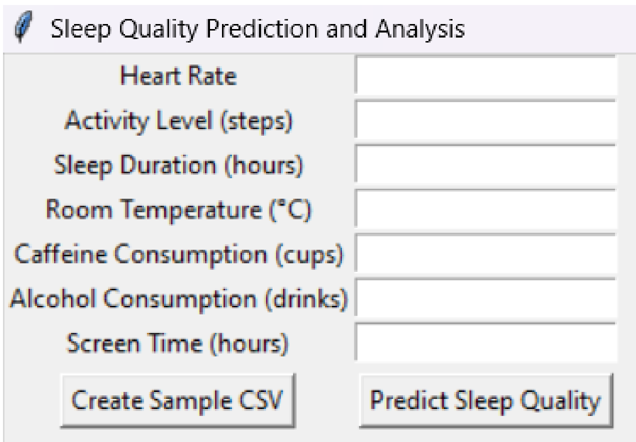


Fig I. The images is showing the user input of their data for sleep quality.

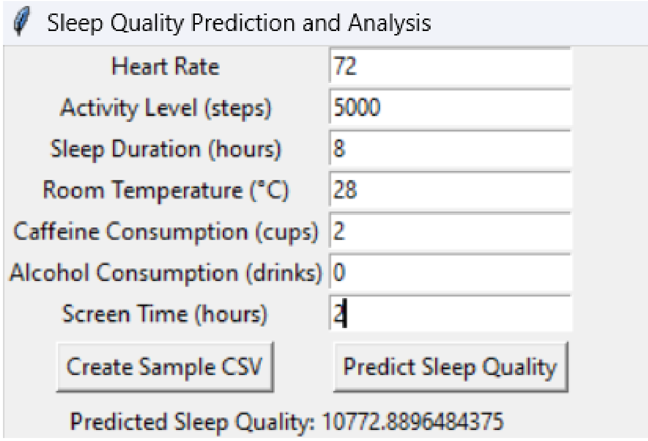


Fig II. Inputting user data and after that it predict the sleep quality .

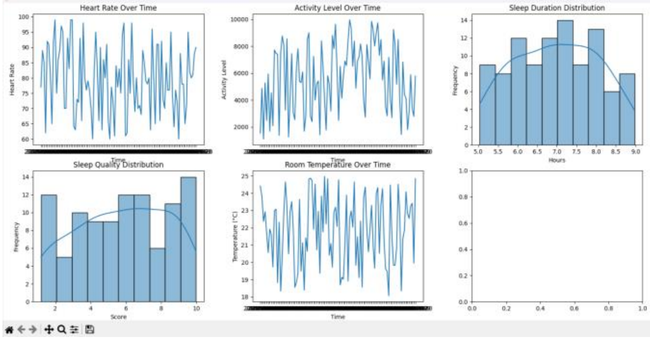


Fig III. After predicting it analysis the data and make the graph format of it so that the user can understand easily through this graph.

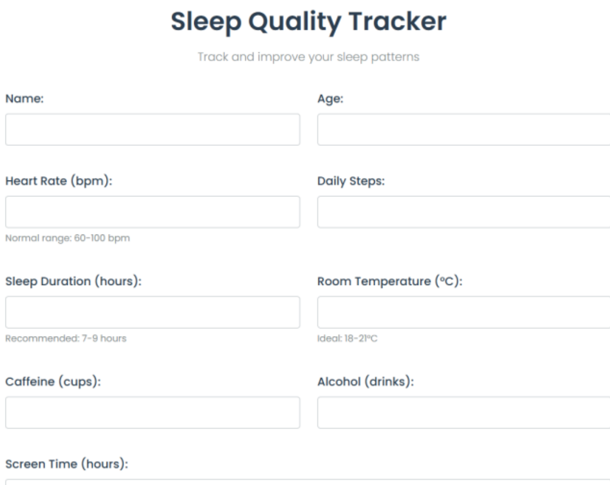
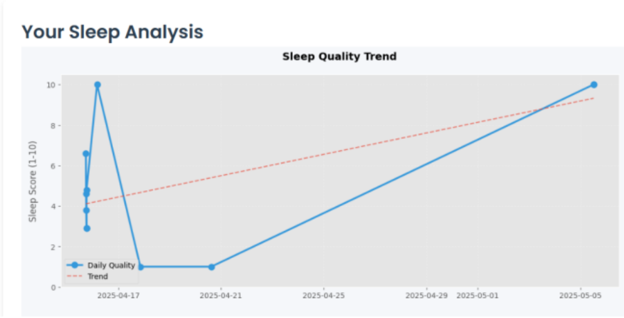


Fig IV. An first page of the site showing the information should be there to insert



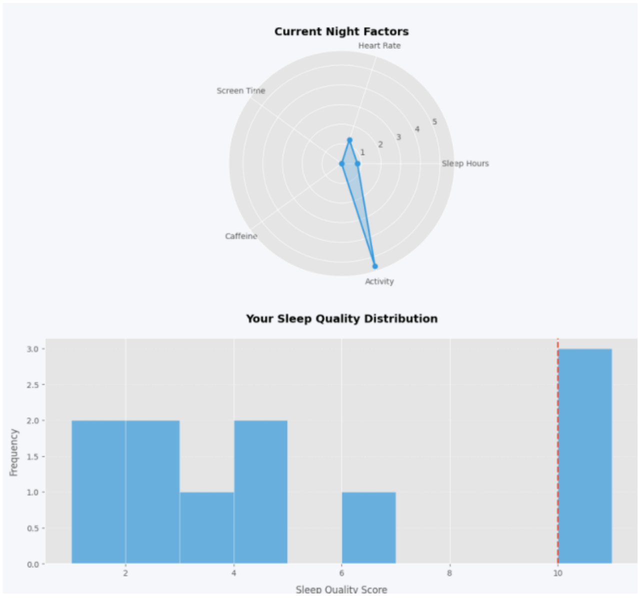


Fig V. Graph analysis of the sleep data also it is 0.82 median accurate

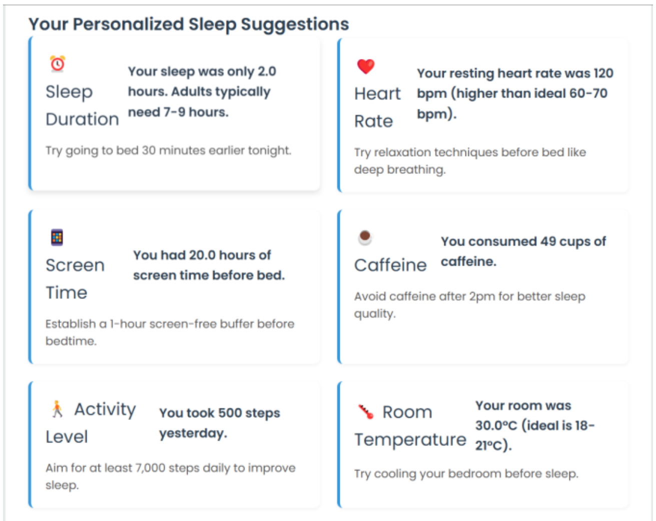


Fig VI. This is Personalised Sleep Suggestion which is required if there any problem or help needed to avoid the harmful things.

5.5 Formula

Neural Network Prediction Formula

The general form of the neural network model used to predict sleep quality is:

$$\hat{y} = f(x) = W_3 \cdot \sigma(W_2 \cdot \sigma(W_1 \cdot x + b_1) + b_2) + b_3$$

Where:
- x is the input feature vector consisting of: heartrate, activity level, sleep duration, room temperature, caffeine consumption, alcohol consumption, screen time
- W_i, b_i are the weight matrices and bias vectors for each layer
- σ is the activation function (REL is used in the hidden layers)
- $\hat{y}$ is the predicted sleep quality score (continuous value between 1 and 10)

Loss Function

The model is trained by minimizing the Mean Squared Error (MSE), which quantifies the difference between the actual and predicted sleep quality scores:

$$MSE = (1/n) \sum(i=1 \text{ to } n) (y_i - \hat{y}_i)^2$$

Where:
- y_i is the actual sleep quality score for sample i
- $\hat{y}_i$ is the predicted sleep quality score
- n is the total number of samples

This approach allows the model to learn non-linear relationships between the physiological and lifestyle features and the corresponding sleep quality outcomes.

VI. RESULTS

The neural network model achieves a mean squared error (MSE) indicating its predictive accuracy. The model's predictions are compared with actual sleep quality ratings, demonstrating its reliability.

VII. DATA VISUALIZATION

The following visualization provides insights into the data and model predictions:

import matplotlib.pyplot as plt import seaborn as sns # Plot sleep quality distribution plt.figure(figsize=(12, 6)) sns.histplot(data['sleep_quality'], bins=10, kde=True) plt.title('Sleep Quality Distribution') plt.xlabel('Score') plt.ylabel('Frequency') plt.show()

VIII. DISCUSSION

Our findings suggest that AI can accurately predict sleep quality based on fitness data and can provide personalized recommendations for improving sleep habits. This study also highlighted the potential of wearable fitness devices for monitoring and enhancing sleep quality.

IX. COMPARISON WITH PREVIOUS RESEARCH

Our results align with those of previous studies that linked various factors to sleep quality. However, this study extends existing research by integrating multiple factors and leveraging AI for predictive analysis.

X. IMPLICATIONS

These findings have significant implications for individuals seeking to improve sleep quality. By using wearable fitness devices and AI, individuals can gain personalized insights and recommendations to enhance their sleep habits.

XI. LIMITATIONS

This study has several limitations that should be considered when interpreting the results. First, the dataset was limited to a relatively small group of participants (100 individuals), which might not represent the diversity of the general population. Larger and more varied datasets could improve the generalizability of the model.

Second, while wearable devices provide convenient data collection, some metrics, such as caffeine or alcohol intake, are self-reported and may be prone to inaccuracy or bias. Incorporating objective sensors or integrating verified external data sources can enhance the accuracy of predictions.

Third, environmental and psychological variables such as stress levels, noise exposure, and emotional well-being were not included in this model. These factors can significantly impact sleep but are harder to reliably quantify using wearable data alone.

Finally, the model uses a regression-based approach. Exploring other techniques, such as classification models or hybrid deep learning frameworks, may yield even more accurate or insightful outcomes in future iterations.

XII. FUTURE SCOPE

To enhance the model and broaden the study’s applicability, future work may include:

Larger and More Diverse Datasets: Incorporating data from users across different age groups, lifestyles, and geographic locations will improve model generalizability.

Additional Features: Including stress levels, emotional state, ambient light/noise levels, and sleep stage segmentation could provide richer input.

Real-Time Applications: Developing a mobile or smartwatch app that processes data in real time and delivers personalized sleep advice.

Comparison with Other ML Models: Benchmarking performance against models like Random Forest, XGBoost, and Support Vector Machines.

Explainable AI (XAI): Using interpretability tools like SHAP or LIME to help users understand how specific behaviors impact their sleep quality predictions.

XIII. CONCLUSION

This project successfully demonstrated that artificial intelligence, when integrated with data from wearable fitness devices, can be used to predict sleep quality with considerable accuracy. The developed model processes a variety of behavioural and physiological inputs to generate sleep quality predictions, offering valuable insights into personal sleep habits.

The outcomes highlight the practical potential of AI-powered health-monitoring systems, especially for proactive wellness and lifestyle management. As wearable technologies become more widespread, the ability to provide personalized sleep recommendations using AI will become increasingly accessible.

In the future, expanding the dataset, refining prediction models, and integrating additional features, such as stress levels or mood tracking, could significantly improve the performance and scope of such systems. Ultimately, this project paves the way for smarter and more individualized healthcare tools that help users make informed decisions for a healthier lifestyle.

XIV. References

[1]. Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index (PSQI): A new instrument for psychiatric practice and research. Psychiatry Research, 28(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

[2]. Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., ... & Adams Hillard, P. J. (2015). National Sleep Foundation's Sleep Time Duration Recommendations: Methodology and Results Summary. Sleep Health, 1(1), 40–43.

[3]. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

[4]. Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1412.6980

[5]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

[6]. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[7]. National Sleep Foundation. (2020). Sleep Health Index. Retrieved from https://www.sleepfoundation.org/professionals/sleep-health-index

[8]. American Academy of Sleep Medicine. (2020). The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.6. Darien, IL: American Academy of Sleep Medicine.