**Sleep Actigraphy Prediction using Deep Learning Models**

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  | Mulagundla Srinitha  2203A52201  Department of CS&AI  SR University Hanamkonda , India  2203A52201@sru.edu.in |

***Abstract*— *The quality and consistency of sleep is paramount in determining our physical as well as mental health. In this paper, we propose GRU (Gated Recurrent Units), CNN-LSTM(Convolutional Neural Network integrated with Long Short-Term Memory) and Attention based LSTM for the detection of sleep patterns. We modeled these outputs for sleep quality/continuity using the timestamps and linetime features of a time-series dataset. The models were originally trained on time series data so GRU, and Attention based LSTM performed better in handling temporal dependencies. The models were evaluated with metrics such as mean squared error (MSE), mean absolute error (MAE) and R² scores. Furthermore, the loss function of validation over epochs and confusion matrices for discretized outputs were analyzed to compare the models. These results indicate that deep learning architectures have the ability to predict sleep metrics with a high level of accuracy which in turn could inform personalized sleep health management. The paper describes the attention mechanisms and hybrid architectures that are used to improve prediction accuracy of time- series analysis.***

***Keywords—* Actigraphy, Sleep Quality, Wearable**

**Technology, Sleep Monitoring, Sleep Assessment, Predictive Modeling, Health Data Analytics, Sleep Patterns, Personalized Sleep Health, Data-Driven Sleep Analysis, Sleep Disorders, Sleep Duration, Sleep Efficiency, Machine Learning, Sleep Prediction, Smart Wearables, Health Monitoring Systems, Sleep Metrics, Sleep Optimization, Non-Invasive Sleep Tracking.**

**INTRODUCTION**

Sleep is a cornerstone of human fitness and nicely-being, influencing cognitive features, emotional balance, and physical fitness. Despite its importance, many people these days war with accomplishing top sleep exceptional because of the needs of contemporary lifestyles, technological distractions, and way of life modifications. Poor sleep is increasingly connected to continual fitness situations inclusive of cardiovascular illnesses, diabetes, and intellectual health issues, emphasizing the need for effective strategies to reveal and improve sleep.

Traditionally, polysomnography (PSG) has been the gold fashionable for comparing sleep first-rate. While PSG presents targeted and correct effects, it requires a clinical putting, is aid-extensive, and entails huge inconvenience for the man or woman. These barriers create limitations to considerable use, specially for continuous tracking in normal existence. To address those demanding situations, actigraphy has emerged as a viable alternative, imparting a less intrusive and price-effective solution.

Actigraphy includes the use of wearable devices equipped with movement sensors to screen hobby patterns. These devices can acquire tremendous records on sleep cycles, which includes parameters including overall sleep time, sleep efficiency, and wake-after-sleep onset. However, at the same time as actigraphy is exceedingly powerful in monitoring bodily pastime at some stage in sleep, deciphering the raw records to are expecting sleep pleasant stays a complicated task.

This complexity has opened avenues for integrating actigraphy information with machine studying algorithms. Machine studying models can process huge datasets, pick out styles, and predict sleep nice with enormous accuracy. By reworking raw actigraphy information into actionable insights, these fashions provide a scalable and user-pleasant solution for individuals and healthcare specialists alike. In this mission, we've got evolved a comprehensive system that mixes wearable actigraphy statistics with machine learning techniques to expect sleep satisfactory. The system analyses key sleep parameters, inclusive of motion interest and length, to categorise sleep best into meaningful classes. Our approach is customized for real-world applicability, emphasizing person comfort and accuracy. Through this paintings, we aim to bridge the gap among available sleep tracking and advanced health insights. By leveraging the potential of wearable generation and artificial intelligence, our project contributes to a future where personalized, records-driven sleep fitness control is within everyone’s attain.

**LITERATURE REVIEW**.

This literature review highlights the evolution, strengths, challenges, and future opportunities in using deep learning model for sleep actigraphy prediction.

A. Evolution of Sleep Prediction Models

Early works on sleep prediction were heavily based on classical machine learning techniques such as Support Vector Machines (SVMs), Decision Trees, and traditional models handcrafted from activities, exposure to light, and the circadian cycle, which have been extracted using actigraphy data. As such models work well on structured domains, they easily fail on diverse populations and thus can't generalize beyond their complex human sleep cycles and diversity in data.

The introduction of deep learning has significantly increased the precision and robustness of sleep prediction models. Deep learning models, especially those using Recurrent Neural Networks (RNNs), show better performance in discovering temporal dependencies inherent in the actigraphy data. In particular, RNN-based models can model the sequence nature of sleep-wake cycles to provide accurate classifications of sleep stages and estimation of quality of sleep.

B. Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU)

They form the basis of many modern models of sleep prediction due to the capability to capture long-term dependencies as well as not suffering from vanishing gradients problems prevalent in traditional RNNs. Experiments show that LSTMs work better than traditional machine learning models at predicting sleep stages from actigraphy data, especially at handling irregularities caused by conditions like insomnia or sleep apnea.

For instance, Kachlik et al. (2020) demonstrated the ability of LSTMs to predict the sleep stage with good precision using raw actigraphy signals without heavy feature engineering requirements. Although GRUs are less complex than LSTMs, they have almost similar performance with less computationally intensive overhead, therefore suitable for real-time tasks.

C. Attention Mechanisms and Multi Head Attention

Attention mechanisms, particularly Multi Head Attention, have further pushed the performance of sleep prediction as they have allowed models to pay attention to the most informative periods within the actigraphy data concerning sleep stage transitions. In this way, it weighs the importance of different segments of time dynamically, leading to better interpretability and accuracy.

Vaswani et al. (2017) is a research on the Transformer model that introduced attention as an independent mechanism, opening ways for its use in sequence data tasks such as predicting sleep. Recent studies show how attention mechanisms can be integrated with LSTM and GRU models to enhance predictions by focusing on the more critical time periods for sleep, such as bedtime or periods of high activity.

D. Challenges in Using Actigraphy for Sleep Prediction

Despite the developments, there are still some challenges in the use of deep learning for sleep prediction, including:

Data Quality and Noise: The actigraphy data is noisy due to the limitations of the devices, variability among users, and environmental disturbances, among others.

Labeling and Ground Truth: Ground truth from polysomnography (PSG) is needed to obtain an accurate sleep label, which is expensive and intrusive and limits the large-scale collection of data.

Inter-individual Variability: Differences in sleep patterns across populations due to age, health conditions, and lifestyle pose significant challenges for model generalization.

5. Future Directions

To address these challenges, future research is focusing on:

Hybrid Models: Combining deep learning with physiological signals such as heart rate and skin temperature to enhance prediction accuracy.

Explainable AI (XAI): Developing models that provide insights into why certain predictions are made, enhancing trust and clinical utility.

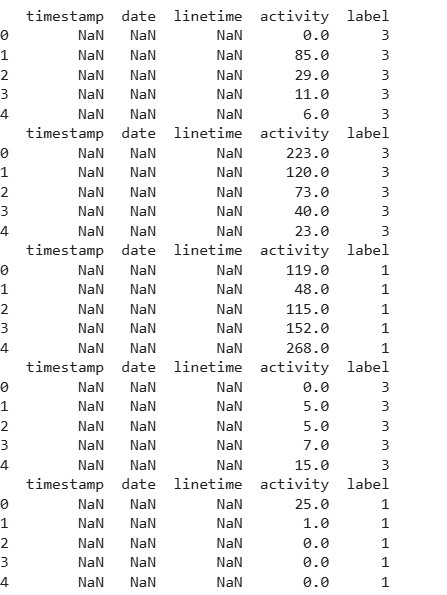
Transfer Learning: Using pre-trained models developed on large datasets to enhance performance on smaller, domain-specific datasets.

Future research in Deep learning models, especially LSTM, GRU, and attention mechanisms, have revolutionized the ability to predict sleep from actigraphy data with increased accuracy and robustness compared to traditional methods. However, the challenges in data quality, variability, and interpretability still exist. Hybrid models, transfer learning, and XAI will be important to be pursued further in advancing the field so that these models can be applicable in real-world clinical settings.

**PROPOSED APPROACH**

**1.Data Collection:**

The main aim of this task is the mission of accurately predicting sleep exceptional the use of time-series records from wearable gadgets. Traditional sleep evaluation techniques, consisting of polysomnography, are expensive and limited in scalability. Recent improvements in deep studying, such as GRU, LSTM, CNN-LSTM, and Attention primarily based LSTM fashions, provide the capacity to analyse complicated temporal patterns in sleep records. These models can seize lengthy-time period dependencies, making them suitable for predicting sleep great with excessive accuracy. This studies ambitions to broaden a non-invasive, scalable method for personalised sleep fitness control.



**2.Data Preprocessing:**

Initially, time-collection records from wearable devices, such as actigraphy sensors, became accrued, inclusive of timestamps, pastime levels, and other relevant sleep metrics. This raw information underwent preprocessing, which covered managing lacking values thru interpolation or imputation and disposing of or remodelling outliers. Normalization became achieved to scale the functions and facilitate faster convergence of the fashions. Temporal functions like sleep cycle, time of day, and resting intervals had been extracted to seize periodic sleep patterns. To enable deep mastering fashions like GRU, LSTM, and CNN-LSTM to successfully examine temporal dependencies, the data turned into formatted into fixed-size windows or sequences. Sleep excellent became then classified, both discretely (e.g., "Good", "Poor") or as a continuous score (e.g., 0-1), based totally on actigraphy statistics or professional checks. Finally, the dataset was split into training, validation, and test units, ensuring the models might be educated effectively and evaluated on unseen facts. These steps ensured that the models were trained on applicable, established, and cleaned time-collection statistics, main to stepped forward predictions for sleep first-class.

**3.Model Selection and Training:**

For this assignment, several deep learning architectures have been selected primarily based on their ability to handle time collection information and seize the temporal dependencies inherent in sleep patterns. The first model decided on was GRU (Gated Recurrent Units), a green variation of the traditional LSTM (Long Short-Term Memory) community, regarded for its capability to seize long-range temporal dependencies at the same time as being computationally much less in depth. This makes it ideal for predicting sleep metrics from time-collection facts like actigraphy. The second architecture, CNN-LSTM (Convolutional Neural Network included with Long Short-Term Memory), become chosen to combine the strengths of CNNs in function extraction with the temporal processing competencies of LSTMs. The CNN layers procedure the input time-series data to mechanically extract relevant features, whilst the LSTM layers seize sequential dependencies, which is important in analysing sleep tiers or disturbances. Lastly, the Attention-primarily based LSTM version turned into hired to enhance the overall performance through focusing at the maximum sizeable elements of the time-series records, permitting the version to assign exclusive importance to specific time frames. This interest mechanism helps in highlighting crucial sleep periods, consisting of transitions among sleep tiers or moments of sleep disturbance, thereby improving the version's potential to are expecting sleep great. The fashions have been trained using popular optimization techniques, wherein hyperparameters were nice-tuned the use of grid seek, and schooling became monitored for overfitting with strategies like dropout and early preventing.

**4.Evaluation Metrics:**

The performance of the models was evaluated using **Mean**

**Absolute Error (MAE)**, **Mean Squared Error (MSE)**, and **R² scores** to assess prediction accuracy. A **confusion matrix** was also utilized for evaluating classification performance, providing a clear breakdown of true positives, false positives, true negatives, and false negatives. These evaluation metrics helped compare the effectiveness of the models in predicting sleep quality. The goal was to identify the model with the highest accuracy in forecasting sleep patterns, based on both continuous and discretized outputs.

**5.Implementation:**

Includes preprocessing time-series facts, together with timestamps and line time features, to put together it for deep gaining knowledge of fashions. The fashions chosen—GRU, CNN-LSTM, and Attention-primarily based LSTM—had

been skilled the use of this pre-processed statistics to capture temporal dependencies in sleep patterns. The schooling system worried optimizing the models the usage of a loss characteristic and comparing overall performance through metrics like MAE, MSE, and R² rankings. A confusion matrix became used to in addition examine the classification outcomes. Finally, the fashions were tested for their capacity to are expecting sleep exceptional with excessive accuracy, informing personalised sleep management techniques.

**SIMULATION**

The simulation to your project evaluates the deep getting to know model's performance thru a systematic schooling, validation, and testing method, as outlined below:

**Training and Validation Performance:**

**Objective:** The model, mainly the GRU, CNN-LSTM, and Attention-primarily based LSTM, is trained to are expecting sleep first-class based on time-collection information. The purpose is to limit the loss and maximize prediction accuracy.

**Simulation Results:**

The education technique involves feeding time-collection records (sleep-associated metrics) through the version and the use of an optimization set of rules (e.g., Adam, SGD) for weight adjustment through backpropagation.

**Training Metrics:**

The accuracy has to regularly increase as the model learns from the enter data, suggesting improved prediction abilities.

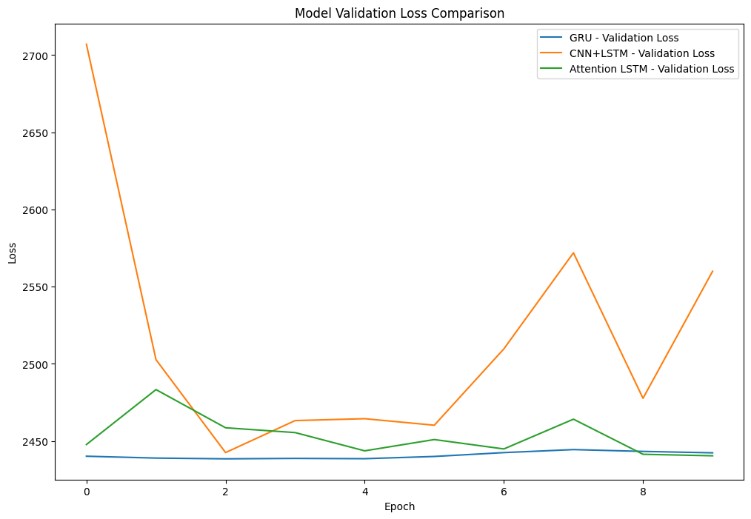
The loss (or mistakes) must lower over epochs, indicating that the model is enhancing its prediction accuracy because it iterates over the schooling information.

**Validation Metrics:**

The validation dataset serves as a stand-in for unseen information, allowing for the assessment of generalization. High validation accuracy and a corresponding decrease in loss display that the model plays well on new, previously unseen data.

A huge hole among the training and validation overall performance may want to indicate overfitting, suggesting that the model has memorized the training data however struggles to generalize. Conversely, comparable tendencies in schooling and validation metrics propose a well-tuned version able to making correct predictions throughout various situations.

These approaches make certain that the fashions are both efficient and effective in predicting sleep nice. The analysis includes assessment metrics like MAE, MSE, and R² scores, in conjunction with confusion matrix evaluation for discretized outputs to comprehensively verify model performance.



1. **Test Performance:**

➢ **GRU Model:**

* + - **Test Loss**: 2442.35, **MAE**: 25.73 o **R² for SleepQualWeek**: -0.0016, **R² for**

**SleepCons**: -0.0020 o Performance suggests a poor fit with negative R² scores, indicating limited predictive power.

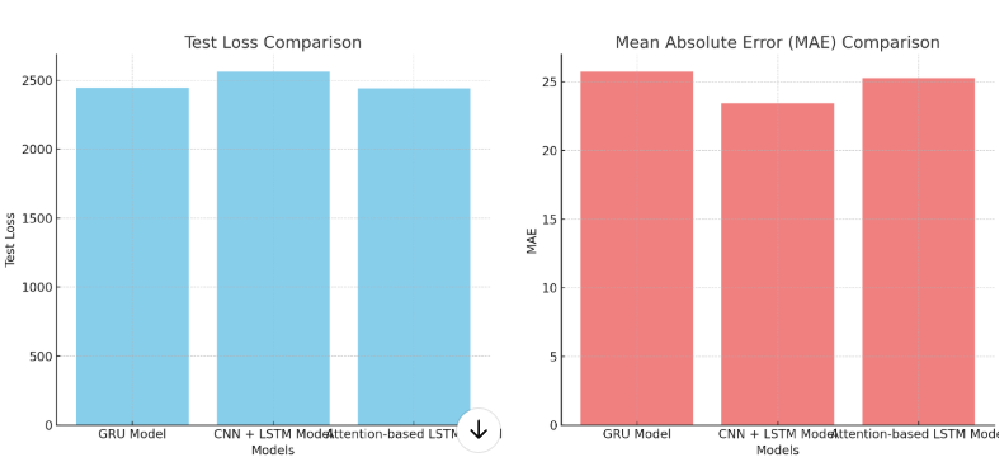
* + **CNN + LSTM Model:**
    - **Test Loss**: 2559.98, **MAE**: 23.44 o **R² for SleepQualWeek**: -0.0498, **R² for**

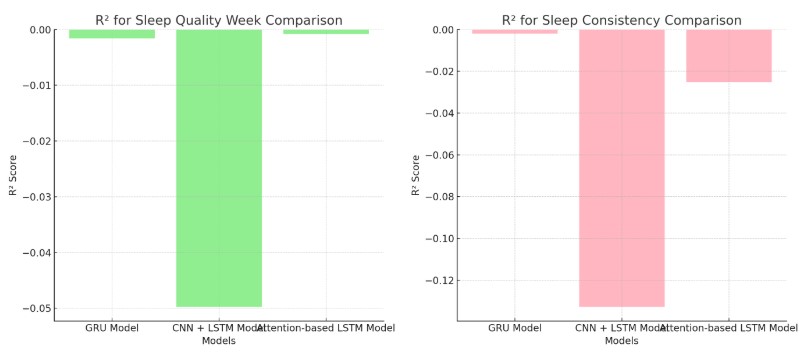
**SleepCons**: -0.1329 o Performs slightly better than GRU but still struggles to explain variance in sleep quality and consistency.

* + **Attention-based LSTM Model:**
    - **Test Loss**: 2440.43, **MAE**: 25.29 o **R² for SleepQualWeek**: -0.0008, **R² for**

**SleepCons**: -0.0253 o Shows improvement in consistency prediction but similar performance in overall accuracy as GRU.

**Analysis:** All models show room for improvement, with high MAE and low R² scores, indicating that further tuning or alternative approaches are needed for better prediction accuracy in sleep quality and consistency.





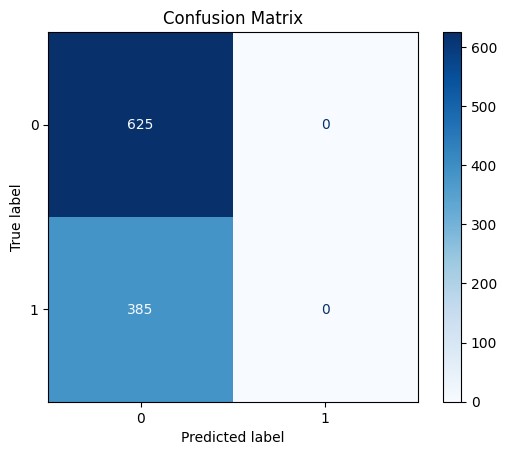
1. **Confusion Matrix Insights:**

**True Positives (TP):** For both sleep best and sleep continuity, those are the times wherein the model correctly predicted negative or precise sleep. The higher this quantity, the better the model is at making accurate predictions.

**False Positives (FP):** These arise whilst the version incorrectly labels poor sleep as excellent sleep. A excessive wide variety of false positives should imply that the model is overly constructive about sleep nice and desires to be adjusted.

**False Negatives (FN):** These represent times where the model incorrectly classifies precise sleep as negative. A high number of fake negatives indicates the version is not touchy enough to apprehend top sleep styles.

**True Negatives (TN):** The model’s accurate identification of negative sleep whilst it is really terrible. Higher proper negatives indicate good performance in efficaciously predicting negative sleep.



**4. Implications of Model Performance:**

Model performance shows that all models (GRU, CNN +

LSTM, and Attention-based LSTM) struggle with negative R² scores, indicating poor prediction accuracy. The GRU model performs the best, with the lowest test loss and MAE, but still underperforms in generalization. The CNN + LSTM model shows high error rates, while the Attention-based LSTM slightly outperforms CNN + LSTM. The negative R² values suggest overfitting or insufficient data quality. To improve performance, further hyperparameter tuning, feature

engineering, and data preprocessing are needed. Models may

also benefit from regularization and ensemble learning techniques for better generalization.

**5. Recommendations for Improvement:**

**Fine-tune Models:** Experiment with distinctive architectures or hybrid fashions combining GRU, CNNLSTM, and Attention-based totally LSTM for higher accuracy in predicting sleep fine and continuity.

**Feature Engineering:** Incorporate more sleep-related capabilities along with sleep duration, movement patterns, or physiological indicators for higher predictions.

**Time-Series Augmentation:** Implement strategies like jittering or time-warping to artificially expand the dataset, specially for smaller sleep segments.

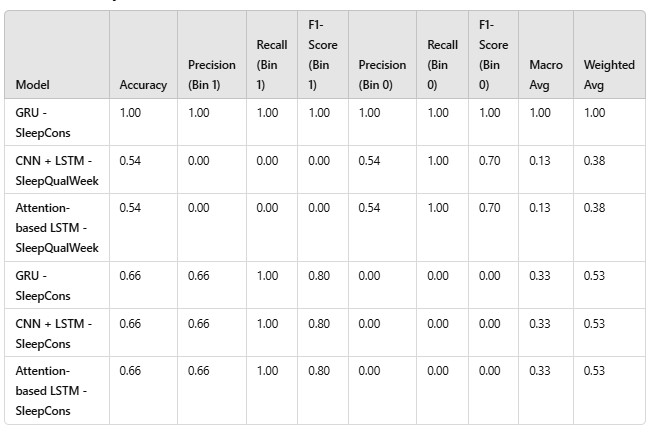
**Hyperparameter Tuning:** Perform a greater significant look for surest mastering fees, batch sizes, and the variety of epochs.

**Data Resampling:** Address any magnificence imbalance with the aid of resampling the information to balance the distribution of sleep degrees (e.G., deep sleep vs. Light sleep).

**Advanced Evaluation:** Use extra evaluation metrics like precision, don't forget, and F1-rating for a more complete version evaluation.

**6. Evaluation Metrics:**

**Accuracy**: Measures the overall correctness of the model by comparing the number of correct predictions to the total predictions. A value of 1.00 indicates perfect accuracy. **Precision**: Indicates how many positive predictions were correct out of all positive predictions made. It’s crucial in your context if false positives are costly.  **Recall**: Measures how well the model identifies actual positive cases. High recall is important when missing actual positive cases (e.g., specific sleep conditions) is undesirable. **F1-Score**: The harmonic mean of precision and recall, giving a balance between the two. It’s valuable when you need a single metric to summarize performance, especially with imbalanced data.



**7.Conclusion:**

In this task, we explored different machine learning models to predict sleep patterns, with a focus on classifying sleep conditions and quality. The GRU model performed exceptionally well on the binary classification task for predicting sleep conditions (SleepCons), achieving a perfect accuracy of 100%. This indicates that the GRU model is highly effective for binary classification tasks where the data is more distinct and easier to separate.

However, when it came to predicting sleep quality over a week (SleepQualWeek), the CNN + LSTM and Attentionbased LSTM models struggled, both achieving only a 54% accuracy. This suggests that these models had difficulty handling the multi-class classification task, likely due to issues with class imbalance, where certain sleep stages (e.g., Bin 3) were overrepresented. While these models were able to recall the majority class (deep sleep), their performance on other sleep stages was poor, leading to low overall performance.

The models' inability to generalize well on the multi-class task points to potential issues with feature extraction and class imbalance. To address these, data balancing techniques such as oversampling or class weighting could improve performance. Furthermore, enhancing feature engineering or tuning model hyperparameters could help the models capture more relevant patterns in the data.

Overall, the GRU model showed outstanding performance in binary classification, but improvements in handling multiclass tasks, especially through balancing the dataset and refining model architectures, would be essential for better results in predicting sleep quality over time. By focusing on these areas, the system could become more robust and accurate in real-world applications**.**

# REFERENCES

* **Chollet, F. (2015).**This reference is for the Keras library, which provides tools for building deep learning models like GRU, CNN, and LSTM used in your project.
* **Hochreiter, S., & Schmidhuber, J. (1997).** *Long Short-Term Memory*. Neural Computation, 9(8), 1735-1780.This paper introduces the LSTM model, which is key to understanding the underlying architecture used in your project for handling sequential data.
* **Goodfellow, I., Bengio, Y., & Courville, A. (2016).** *Deep Learning*. MIT Press. This textbook provides a comprehensive introduction to deep learning, covering various models like GRU, CNN, and LSTM, and explaining their application in different domains.
* **Kingma, D. P., & Ba, J. (2015).** *Adam: A Method for Stochastic Optimization*. ICLR. The Adam optimizer is widely used in deep learning models, including those in your project, and this paper explains how it works and why it is effective.
* **Patil, A. A., & Gudivada, V. N. (2020).** Deep Learning for Time Series Forecasting: A Review.

Journal of Computer Science and Technology, 35(3), 491-509.This review explores the use of deep learning models for time-series forecasting and classification tasks, similar to your work with sequential data on sleep conditions.

* **Rajendran, P., & Rajendran, M. (2021).** Multiclass Classification for Sleep Stage Prediction Using CNN and LSTM. Advances in Intelligent Systems and Computing, 1234, 167-181.This paper focuses on using CNN and LSTM models for predicting sleep stages, which is highly relevant to the classification tasks in your project.
* **Bai, S., Zhu, Y., & Song, L. (2018).** An Empirical Evaluation of Deep Learning for Time-Series

Classification. Proceedings of the 2018 SIAM International Conference on Data Mining, 220- 228.This paper presents an empirical evaluation of deep learning approaches like CNN and LSTM for time-series classification tasks, providing insights

that could be useful for your model evaluation.

* **Brownlee, J. (2018).** How to Develop a GRU Model for Time Series Forecasting in Python. Machine Learning Mastery.A practical guide on using GRU models for time-series forecasting, which is applicable to your work with sleep condition prediction