A Deep Learning approach to detect sleep states

Armando Bringas, Alexis Guerrero

3rd January - 2024

Introduction

Sleep States Detection

Our project proposal rooted in the practical application of a Kaggle competition, aimed at addressing a challenge in pediatric health and neuroscience: the identification of sleep states in children through wrist-worn accelerometer data
[Esper et al.(2023)Esper, Demkin, Hoolbrok, Kotani, Hunt, Leroux, van Hees, Zi

Project implications

The significance of this project lies in its potential to deepen our understanding of sleep and to provide further insights into its importance. For instance, understanding how environmental factors influence sleep, mood, and behavior can aid in formulating personalized strategies tailored to the unique needs of each child.

Literature Review

Models

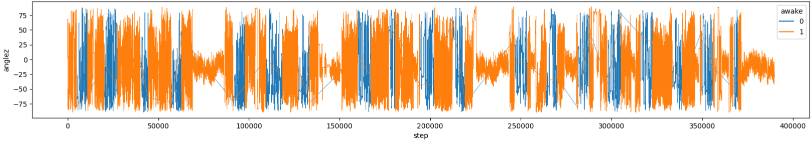
- Random Forest
- Residual Neural Network (RNN)
- Latent Class Analysis (LCA)
- Long Short-Term Memory (LSTM)

Data origin

- Accelerometer data
- Actigraphy and sleep diaries
- Optical plethysmography
- Electrocardiography
- Electroencephalography
- Multi channel polysomnogram
    - Electroculography
    - Electromyography

Dataset

Dataset consists of approximately 500 multi-day recordings from wrist-mounted accelerometers. The accelerometer data in the dataset was processed using R with the GGIR package [Migueles et al.(2019)Migueles, Rowlands, Huber, Sabia, and van Hees]. The recordings are labeled with two event types: 'onset', indicating the start of sleep, and 'wakeup', marking its end. The primary objective is to identify these two events within the accelerometer data series, this primarily represents a binary classification task.



Model - Random Forest

The Random Forest model is initialized with a predetermined number of 100 estimators (n_estimators) and a minimum leaf sample size (min_sample_leaf) set to 300. These initial values are chosen for the purpose of assessing the model's efficiency and will be subject to subsequent adjustments based on its performance. Given the substantial volume of data in this scenario, it is anticipated that the minimum leaf sample size will remain unchanged, whereas the number of estimators will likely need to be increased to optimize performance. An iteration process is set along the Random Forest model, changing the random_state and n_estimators values.

$$\mathcal{X}_{RF} \longrightarrow \boxed{\begin{array}{c}\text{Random Forest}\\ i\text{th iteration}\end{array}} \longrightarrow \mathcal{Y}_{RF_i}(rs_i, Ne((i-1)[\mathrm{mod}3]))$$

Model - LSTM

We are proposing a starting neural network architecture with the following blocks where consists of an LSTM layer of 64 units, ideal for processing sequences by capturing dependencies from prior inputs. This is followed by a Dense layer, the size of which matches the number of classification categories in your problem, in this case for our binary classification problem $n_{\text{classes}} = 2$. The final component is a softmax activation function, applied to convert the output into a probability distribution across the predicted classes.

$$\xrightarrow{\text{input}} \boxed{\text{LSTM}(64)} \rightarrow \boxed{\text{Dense}(n_{\text{classes}})} \xrightarrow{\text{softmax}}$$

Figure: Neural Network for accelerometer data classification

Evaluation Metrics

Average Precision

$$\text{Inputs } \mathcal{X} = (x_1, \ldots, x_N)^\intercal$$
$$\text{Outputs } \mathcal{Y} = (y_1, \ldots, y_N)^\intercal$$
$$\text{Labels } \mathcal{W} = (w_1, \ldots, w_N)^\intercal$$

$$AP = \frac{1}{N} \sum_{i=1}^{N} \mathcal{C}(w_i, y_i)$$

$$\mathcal{C}(w_i, y_i) = \begin{cases} 1 \text{ if } w_i = y_i \forall i \in \mathbb{N}, 1 \leq i \leq N \\ 0 \text{ if } w_i \neq y_i \forall i \in \mathbb{N}, 1 \leq i \leq N \end{cases}$$

$$APF(\mathcal{W}, \mathcal{Y}) = \frac{1}{N} (\omega_1, \ldots, \omega_N)(1, \ldots, 1)^\intercal$$

Evaluation Metrics

Specificity and Recall

$$\text{Score}(x) = \begin{cases} \text{TP} & \text{if } x \text{ matched and } x > \text{thresh.} \\ \text{FP} & \text{if } x \text{ unmatched pred.} \\ \text{FN} & \text{if } x \text{ unmatched truth} \\ \text{TN} & \text{otherwise} \end{cases}$$

$$\text{Specificity} = SPC = \frac{TP}{TP + FN} \qquad \text{Recall} = SN = \frac{TN}{TN + FP}$$

Final evaluation

$$\text{Model evaluation } = \alpha APF + \beta SPC + \gamma SN$$

Results

Due to computing performance implications we perform training and evaluation in a reduced dataset consisting in 35 recording out of the 500 recording provided as well by the Kaggle competition.

Table: Random Forest Confusion Matrix

Actual / Predicted	**Predicted Negative**	**Predicted Positive**
Actual Negative	786677 (31.38%)	56959 (2.27%)
Actual Positive	70514 (2.81%)	1592674 (63.53%)

Table: LSTM Confusion Matrix

Actual / Predicted	**Predicted Negative**	**Predicted Positive**
Actual Negative	813470 (32.45%)	30166 (1.20%)
Actual Positive	178414 (7.12%)	1484774 (59.23%)

Results

Random Forest slightly outperformed LSTM in the Confusion Matrix for Accuracy, Recall, and Specificity metrics. However, concerning model training computational performance, Random Forest took approximately 1 hour to train a single iteration, compared to LSTM, which took approximately 10 minutes to train all defined epochs.

Table: Model Comparison

Metric	**Random Forest**	**LSTM**
Accuracy	0.95	0.92
Recall	0.96	0.90
Specificity	0.96	0.97
Average Precision	0.99	0.99
Partial Evaluation	0.96	0.95

Conclusion

In conclusion, our findings underscore the importance of considering both performance metrics and computational efficiency when selecting a model for sleep state detection. Moreover, our exploration underscores the intricate nature of sleep state detection, compounded by the diverse origins of data stemming from various devices, making data sources a relevant point to take into consideration when selecting a model. Additionally, the need for context-specific solutions and the exploration of advanced architectures indicate exciting avenues for future research in this field, enhancing accessibility to personalized care tailored to the unique characteristics of each patient in the study of sleep.

Bibliography

Nathalia Esper, Maggie Demkin, Ryan Hoolbrok, Yuki Kotani, Larissa Hunt, Andrew Leroux, Vincent van Hees, Vadim Zipunnikov, Kathleen Merikangas, Michael Milham, Alexandre Franco, and Gregory Kiar.
Child mind institute - detect sleep states, 2023.
URL https://kaggle.com/competitions/child-mind-institute-detect-sleep-states.

Jairo H. Migueles, Alex V. Rowlands, Florian Huber, Séverine Sabia, and Vincent T. van Hees.
Ggir: A research community–driven open source r package for generating physical activity and sleep outcomes from multi-day raw accelerometer data.
Journal for the Measurement of Physical Behaviour, 2(3):188–196, 2019. doi: 10.1123/jmpb.2018-0063.