Deep Learning-based classification of Sleep Stages

Razi Mahmood
UC Berkeley

Background of Sleep Stages:

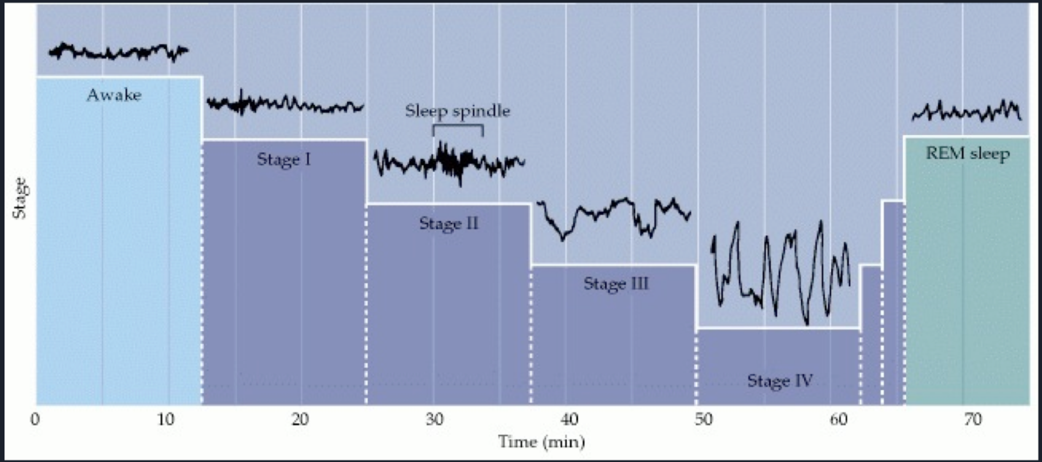
- Sleep is necessary for the body to perform restorative functions such as growing tissues or forming memories and giving rest to brain and body
- During sleep, our body transitions through **different sleep stages** constituting several sleep cycles.
- The stages of the sleep cycle are:
  - **W - Awake**
  - **N1 - Stage 1**
  - **N2 - Stage 2**
  - **N3/N4 - Stage 3/4**
  - **R - REM (Rapid eye movement)**

Chambon, S., Galtier, M. N., Arnal, P. J., Wainrib, G., & Gramfort, A. (2018). A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(4), 758-769.

- The sleep cycle duration, and patterns vary between people as a function of many factors including
  - Age
  - Stress level
  - Professions (athletes versus students)
  - Health conditions
- Studying variability of the length of the stages across sleep cycles & across patients can give a good idea of the expected patterns
- Assess condition of a current patient against the average sleep behaviors
  - Help diagnose sleep issues early on to take corrective measures ahead of time

What is the problem addressed?

- The sleep stages can be seen through **EEG (electro-encephalograph) recordings.**
- These are electrodes placed at various locations on the brain surface that **measure the electrical activity** in the brain and **record the frequency of waves** such as the **alpha wave, beta wave, delta wave**, etc.
- The EEG recordings differ in different sleep stages as below
- By studying the EEG time series, we can automatically **classify the sleep stages through machine learning** approaches
- After classification of stages, we can study **how long each person is spending in each sleep stage** on average



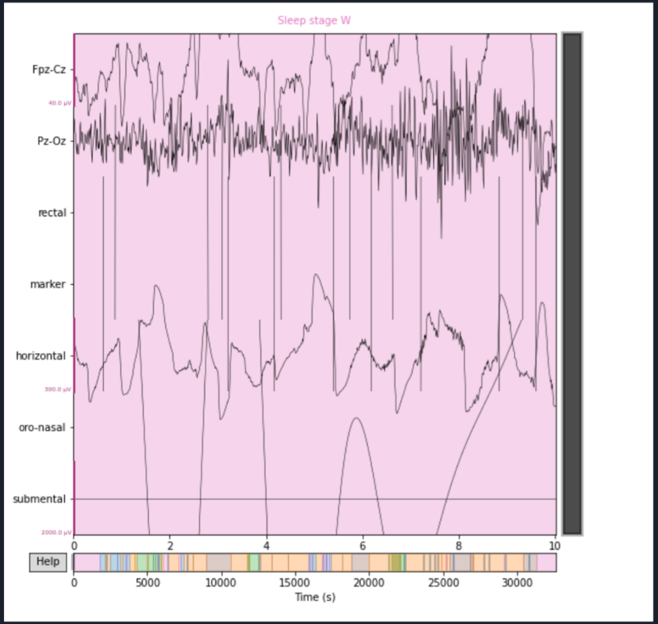
Neuroscience, 2nd edition

Editors: Dale Purves, George J Augustine, David Fitzpatrick, Lawrence C Katz, Anthony-Samuel LaMantia, James O McNamara, and S Mark Williams. Sunderland (MA): Sinauer Associates; 2001.

Why is this a difficult problem?

- Real sleep stage **patterns are very messy**!
- The **wave patterns differ across people** and electrodes (area of the brain tapped)
- Reliable recognition of sleep stages is a difficult problem
- Previous approaches used engineered features through time or frequency domain analysis (e.g. power spectrum, or wavelets) to detect and classify sleep stages
- With **deep learning** algorithms, it has now become possible to **build classifiers** that can be **trained** from **large open source datasets**

Sleep stage waves are messy!



Chambon, S., Galtier, M. N., Arnal, P. J., Wainrib, G., & Gramfort, A. (2018). A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(4), 758-769.

Proposed Creative Approach:

- Adopt a **deep learning approach** to classifying the sleep stages
- Study the classified sleep stage length patterns inferred and compare it to ground truth length patterns recorded
- Although there are several deep learning packages available, I worked off the deep learning architecture and code base proposed in https://github.com/hubertjb/dl-eeg-tutorial
- Made the following changes/experiments:
    - The code didn't work as is because the **architecture was programmed for 7 channel input,** whereas the dataset we worked on (/mne_data/physionet-sleep-data/SC4001E0-PSG.edf) had **only 5 channels**.
        - **Changed the architecture to take 5 channels** instead
    - The deep learning code mashed all data across patients into the training and test datasets
    - Extracted the number of samples per patient and separated the results of deep learning predictions back per patient to reconstruct the time series of sleep stage prediction per patient
    - Analyzed the **intra** and **inter-sleep stage changes** to **record average length** of time for **duration of each sleep stage** per patient
    - Summarized these results across patients to get a result on the variability across patients between the duration of their sleep stages

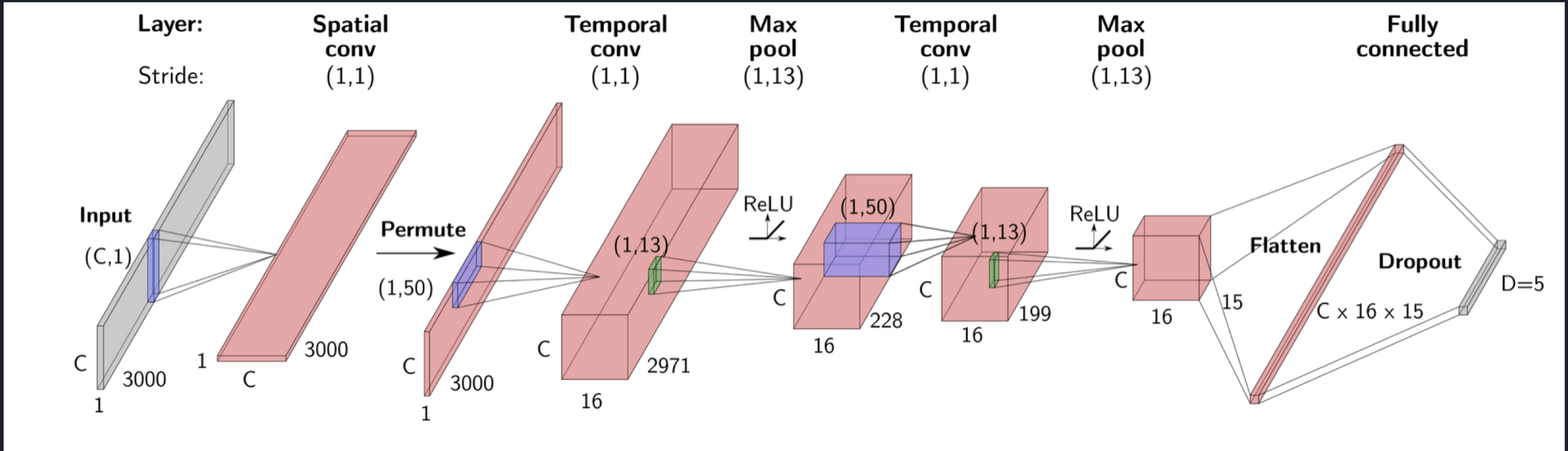
Deep learning approach to sleep stage classification

Architecture chosen: 1 D convolutional neural network (8 layers, 1 drop-out layer)

Input: 30 second window time series per channel

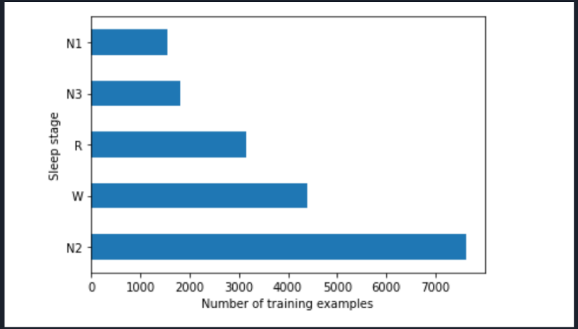
Output: classification layer for one of the 5 sleep classes (W, N1, N2, N3, R)

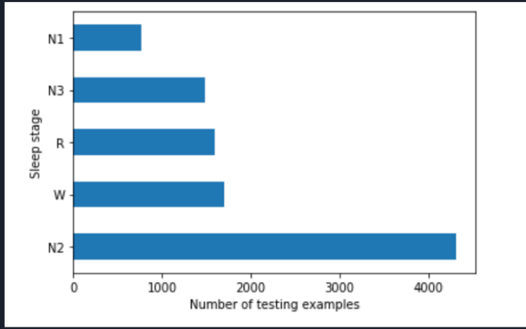
Chambon, S., Galtier, M. N., Arnal, P. J., Wainrib, G., & Gramfort, A. (2018). A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(4), 758-769.



Deep learning training details

Chambon, S., Galtier, M. N., Arnal, P. J., Wainrib, G., & Gramfort, A. (2018). A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(4), 758-769.





Dataset : 30 patients

Training: 20 patients

Testing: 10 patients

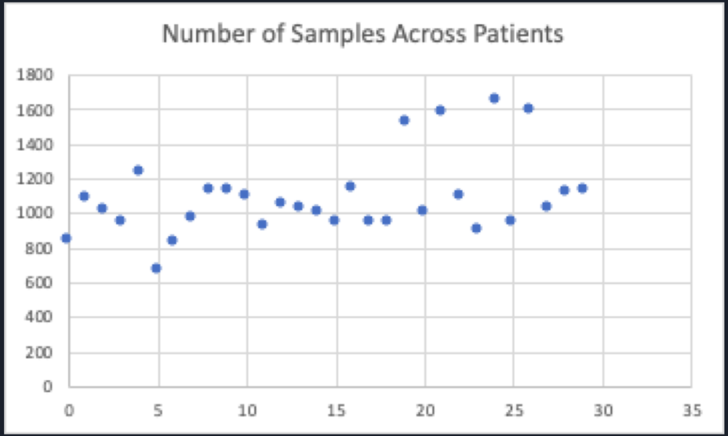
Number of training samples: 18545 (train), 4094 (validate)

(80-20 split for training and validation)

Testing samples: 9850

Number of training epochs: 10

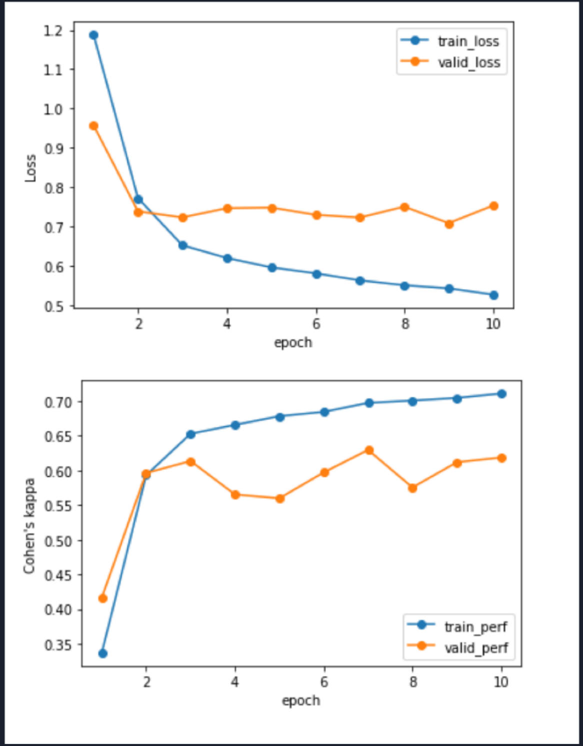
Sample distribution of test samples across patients →



Deep learning training and classification results

- Test balanced accuracy: 0.760
- Test Cohen's kappa: 0.706

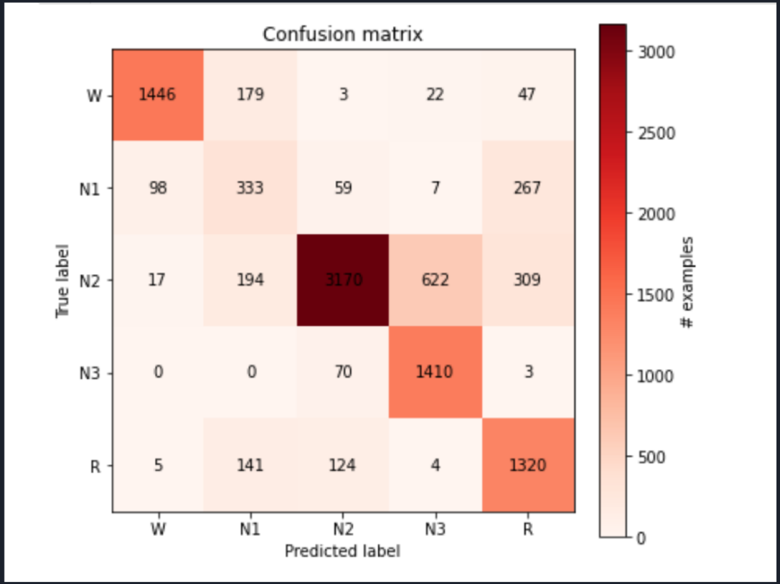
```
epoch     train_loss      valid_loss      train_perf      valid_perf
---------------------------------------------------------------------
1         1.1875          0.9571          0.3370          0.4166
best val loss inf -> 0.9571
2         0.7719          0.7382          0.5932          0.5963
best val loss 0.9571 -> 0.7382
3         0.6520          0.7230          0.6528          0.6136
best val loss 0.7382 -> 0.7230
4         0.6197          0.7462          0.6657          0.5653
5         0.5959          0.7479          0.6783          0.5598
6         0.5808          0.7297          0.6845          0.5972
7         0.5628          0.7228          0.6974          0.6295
best val loss 0.7230 -> 0.7228
8         0.5504          0.7502          0.7007          0.5754
9         0.5425          0.7085          0.7046          0.6119
best val loss 0.7228 -> 0.7085
10        0.5266          0.7529          0.7109          0.6186
```



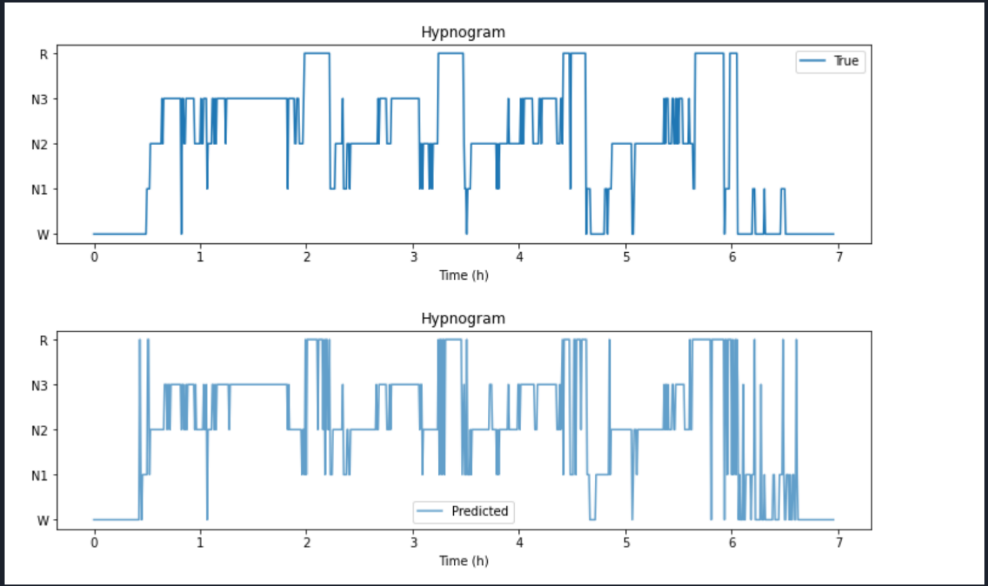
Sleep stage predictions using deep learning

- Sleep stage N1 has more classification errors as seen from confusion matrix

Confusion Matrix of predicted labels



Top: Ground truth sleep stages for one patient with 836 samples



Bottom: Predicted sleep stages for one patient with 836 samples

Study on the variability of sleep stages across patients:

Key Approaches:

- Extract the **number of samples** per patient
- **Separate the results** of deep learning predictions back per patient to **reconstruct time series of sleep stage prediction** per patient
- **Analyze intra and inter-sleep stage changes** to record the **average length of time** for the duration of each sleep stage per patient
- Summarize results across patients to get a result on the variability across patients between the duration of their sleep stages

Study conclusions:

- Although there are errors in deep learning prediction, the trends are roughly similar even if the absolute value ranges are different.
- In general phase W (awake) and R (REM) sleep are the longest indicating normal sleep patterns for these patients.
- Patient 6 appears to be an outlier

Results of sleep stage analysis:

Variation in average sleep stage durations across patients

Average across all 10 patients:

(ground truth/predicted)

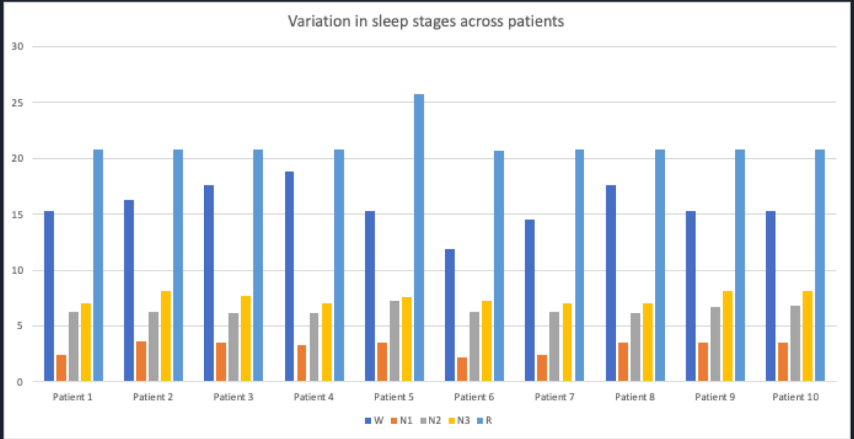
W - 15.82/7.88

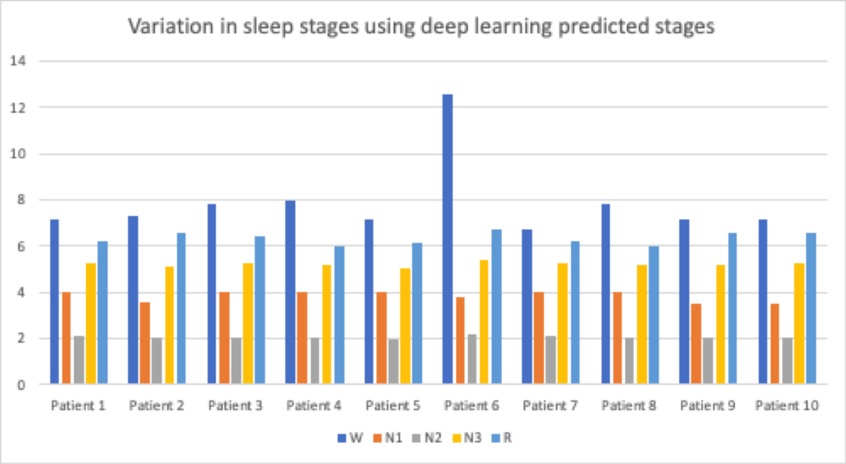
N1 - 3.18/3.85

N2 - 6.43/2.06

N3 - 7.53/5.22

R - 21.31/6.34





Summary

- Study of sleep stages is an interesting problem to explore the value of neurotechnology
- Deep learning can give reasonable approach to automatic sleep stage classification
- Analysis of the average sleep stage patterns can reveal anomalies in patient sleep cycles pointing to potential health problems.
- Adopted a deep learning algorithms and added post analysis enhancements to derive these statistics about average length of sleep stages across patients

Bibliography:

Yildirim, O., Baloglu, U. B., & Acharya, U. R. (2019). A Deep Learning Model for Automated Sleep Stages Classification Using PSG Signals. *International journal of environmental research and public health*, *16*(4), 599. https://doi.org/10.3390/ijerph16040599

Nedelec M, Aloulou A, Duforez F, Meyer T, Dupont G. The Variability of Sleep Among Elite Athletes. *Sports Med Open*. 2018;4(1):34. Published 2018 Jul 27. doi:10.1186/s40798-018-0151-2
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6063976/

Malafeev, A., Laptev, D., Bauer, S., Omlin, X., Wierzbicka, A., Wichniak, A., Jernajczyk, W., Riener, R., Buhmann, J., & Achermann, P. (2018). Automatic human sleep stage scoring using Deep Neural Networks. Frontiers in Neuroscience, 12. https://doi.org/10.3389/fnins.2018.00781

Combrisson, E., Vallat, R., Eichenlaub, J.-B., O'Reilly, C., Lajnef, T., Guillot, A., Ruby, P. M., & Jerbi, K. (2017). Sleep: An open-source python software for visualization, analysis, and staging of Sleep Data. Frontiers in Neuroinformatics, 11. https://doi.org/10.3389/fninf.2017.00060

Ranjan, R., Arya, R., Fernandes, S. L., Sravya, E., & Jain, V. (2018). A fuzzy neural network approach for automatic K-complex detection in sleep EEG signal. Pattern Recognition Letters, 115, 74–83. https://doi.org/10.1016/j.patrec.2018.01.001

Banville, H., Chehab, O., Hyvärinen, A., Engemann, D.-A., & Gramfort, A. (2021). Uncovering the structure of clinical EEG signals with self-supervised learning. Journal of Neural Engineering, 18(4), 046020. https://doi.org/10.1088/1741-2552/abca18

Cox, R., & Fell, J. (2020). Analyzing human sleep EEG: A methodological primer with code implementation. Sleep Medicine Reviews, 54, 101353. https://doi.org/10.1016/j.smrv.2020.101353

Vasko, R., Detka, C., Monahan, J., Reynolds III, C., & Kupper, D. (1996). Muscle artifacts in the sleep EEG: Automated detection and effect on all-night EEG Power Spectra. Journal of Sleep Research, 5(3), 155–164. https://doi.org/10.1046/j.1365-2869.1996.00009.x

Lotte F. (2014) A Tutorial on EEG Signal-processing Techniques for Mental-state Recognition in Brain–Computer Interfaces. In: Miranda E., Castet J. (eds) Guide to Brain-Computer Music Interfacing. Springer, London. https://doi.org/10.1007/978-1-4471-6584-2_7

Banville, H. (n.d.). *HUBERTJB/DL-EEG-tutorial: Hands-on tutorial on deep learning for EEG classification.* GitHub. Retrieved November 20, 2021, from https://github.com/hubertjb/dl-eeg-tutorial.

Sleep stage classification from Polysomnography (PSG) data¶. Sleep stage classification from polysomnography (PSG) data - MNE 0.24.0 documentation. (2021, November 11). Retrieved November 20, 2021, from https://mne.tools/stable/auto_tutorials/clinical/60_sleep.html.