

Using PCA on EEG Data to Distinguish Sleep Stages

Ida Hönigmann

Technical University Vienna, Austria

Email: e12002348@student.tuwien.ac.at

Abstract—

I. INTRODUCTION

II. STUDY OF LITERATURE

first work on pca [9] and [4]
 given paper [6]
 when does pca fail? [11] and [12] (non-linear method)
 book containing sleep phases eeg [2]
 Review Paper on Sleep Stage Classification Methods [1]
 papers trying to solve similar problem [13] and [10] and [7]
 competition using similar data set [3]
 winner of competition [5]

A substantial body of scientific research has been devoted to exploring Principal Component Analysis (PCA). The foundation of this method was laid by Pearson [9] and Hotelling [4].

An introduction to PCA, as well as a good overview on how to derive the formula used to compute the Principal Components (PC) is given by Shlens [11]. Recent applications and variants of PCA are explored by Jolliffe et. al. [6].

Shlens discusses the limitations of PCA, as well as examples in which PCA fails [11], such as the requirement of linearly dependent data. Tenenbaum proposes a non-linear method to combat this problem [12].

Generally speaking the variables must not have third or higher order dependencies between them. In order to reduce a problem dealing with higher order dependencies to a second order one a non-linear transformation of the data can be done beforehand. This method is called kernel PCA [11].

Another method for combating this problem is Independent Component Analysis (ICA) which is discussed by Naik et. al. [8].

In his book Ganong describes typical patterns observed in electroencephalogram (EEG) data of a sleeping person [2]. He explains the EEG patterns associated with rapid eye movement (REM) sleep and non-REM (NREM) sleep, which is further partitioned into four Stages.

The given problem of distinguishing sleep stages given some EEG data has been investigated by use of PCA, as well as neural networks. Some of these works are summarized below.

A review of different methods in the preprocessing, feature extraction and classification is given by Boostani [1].

Tăuțan et. al. [13] compare different methods of dimensionality reduction on EEG data, such as PCA, factor analysis and autoencoders. They conclude that PCA and factor analysis improves the accuracy of the model.

Putilov [10] used PCA to find boundaries between Stage 1, Stage 2 and Stage 3. Changes in the first two PC were related to changes between the Stage 1 and Stage 2, while changes in the fourth PC exhibited a change in sign at the boundary of Stage 2 and Stage 3.

Metzner et. al. [7] try a machine learning approach to try to rediscover the different human-defined-stages. They find that using PCA on the results makes clusters apparent.

The PhysioNet/Computing in Cardiology Challenge 2018 was a competition using a similar data [3]. The goal was to identify arousal during sleep from EEG, EOG, EMG, ECG and SaO2 data given. The winning paper of this competition describes the use of a dense recurrent convolutional neural network (DRCNN) consisting of multiple dense convolutional layers, a bidirectional long-short term memory layer and a softmax output layer [5].

As shown in this section, the utilization of PCA to analyze EEG data has been used with good success.

III. MATHEMATICAL BASICS

IV. PRINCIPAL COMPONENT ANALYSIS

V. SLEEP STAGES AND EEG DATA

VI. DATA AND ALGORITHM

- 1) subdivide eeg signals in the temporal domain
- 2) apply fft transforming into frequency domain
- 3) pca
- 4) achieve dimensionality reduction
- 5) classification of sleep stages
- 6) visualisation

VII. RESULTS

VIII. CONCLUSION

REFERENCES

- [1] Reza Boostani, Foroozan Karimzadeh, and Mohammad Nami. A comparative review on sleep stage classification methods in patients and healthy individuals. *Computer Methods and Programs in Biomedicine*, 140:77 – 91, 2017. Cited by: 212.

- [2] William F. Ganong. *Review of medical physiology*. Appleton & Lange, Stamford, Conn, 18. ed edition, 1997.
- [3] Mohammad M Ghassemi, Benjamin E Moody, Li wei H Lehman, Christopher Song, Qiao Li, Haoqi Sun, Roger G Mark, M Brandon Westover, and Gari D Clifford. You snooze, you win: the physiology/computing in cardiology challenge 2018. *2018 Computing in Cardiology Conference (CinC)*, pages 1–4, 2018.
- [4] Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417, 1933.
- [5] Matthew Howe-Patterson, Bahareh Pourbabae, and Frederic Benard. Automated detection of sleep arousals from polysomnography data using a dense convolutional neural network. In *2018 Computing in Cardiology Conference (CinC)*, volume 45, pages 1–4. IEEE, 2018.
- [6] I. T. Jolliffe and J. Cadima. Principal component analysis: a review and recent developments. *Royal Society*, 374(2065), 2016.
- [7] Claus Metzner, Achim Schilling, Maximilian Traxdorf, Holger Schulze, Konstantin Tziridis, and Patrick Krauss. Extracting continuous sleep depth from eeg data without machine learning. *Neurobiology of Sleep and Circadian Rhythms*, 14, 2023. All Open Access, Gold Open Access, Green Open Access.
- [8] Ganesh R Naik and Dinesh K Kumar. An overview of independent component analysis and its applications. *Informatica*, 35(1), 2011.
- [9] Karl Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 2(11):559–572, 1901.
- [10] Arcady A. Putilov. Principal component analysis of the eeg spectrum can provide yes-or-no criteria for demarcation of boundaries between nrem sleep stages. *Sleep Science*, 8(1):16–23, 2015.
- [11] Jonathon Shlens. A tutorial on principal component analysis. 2014.
- [12] J.B. Tenenbaum, V. De Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319 – 2323, 2000. Cited by: 10812.
- [13] Alexandra-Maria Tăuțan, Alessandro C. Rossi, Ruben de Francisco, and Bogdan Ionescu. Dimensionality reduction for eeg-based sleep stage detection: comparison of autoencoders, principal component analysis and factor analysis. *Biomedical Engineering / Biomedizinische Technik*, 66(2):125–136, 2021.