Sleep Stage Classification

1st Lozzi Daniele
Dept. Computer Science
University of L'Aquila
L'Aquila, Italy
daniele.lozzi@graduate.univaq.it

***Abstract*—This project aims to correctly classify the epochs (30s of signal) of the EEG signal acquired on 74 subjects during the night while they sleep. The tested models show good accuracy, although they are simple and light.**

***Index Terms*—EEG, Sleep Stage Classification, PSD, Machine Learning**

I. Introduction

Sleep is an important physiological stage of daily routine for humans and animals, based on the 24 h cycle mainly directed by the Earth rotation that determines the apparent movement of the Sun. The role of sleep is not well known although many theories have been developed. One of them, homeostatic theory, explains sleep pressure as a need for sleep to remove metabolic waste products and consolidate synaptic connections built during waking activity. Some pathologies that determine a deterioration of sleep stages can have a very negative impact on the individual's social, personal, work, and physical life, and it is necessary to diagnose sleep pathologies in advance to intervene in therapy as soon as possible. Rare pathologies lead to the total disappearance of sleep, which inevitably determines the death of the individual after a few weeks. Unstable sleep determines a moderate-to-severe cognitive deficit depending on personal physiology and the duration of the sleep deficit. [5]

II. Sleep Stage

The existence of Sleep Stage has been well known since the 1950s, when Aserinsky and Kleitman [2] discovered the presence of rapid eye movements during sleep, and from this characteristic they called it REM, that is, Rapid Eye Movements. At the end of the 1960s, after the discovery of REM and NREM sleep and of the concept of cyclicity of these two phases within sleep, the need to classify the EEG variations that occurred macroscopically during sleep in a standard way. These EEG variations were identified with the concept of sleep architecture. In 1968 Rechtschaffen and Kales [9], based on the analysis of electroencephalographic [10], electromyographic and electrooculographic parameters, classified sleep into 5 stages:

- 4 NREM stages (stage 1; stage 2; stage 3; stage 4)
- 1 stage REM.

The presence of this regular alternation of non-REM and REM phases consists of cycles of similar duration to each other.

Sleep Stage Classification is classically performed by visually inspecting the EEG track during the night in real-time or offline mode after recording night. This approach is a very time-consuming and tedious procedure. Again, from the literature is known that each sleep stage has a specific predominance of band frequency respect to others. Many models have been developed to classify sleep epochs [7, 1]. In this work, respecting to the others previous published, we use the Relative Band Power to be independent from the subjective physiological difference in voltage.

III. Materials and Methods

In this work the Sleep-EDF Database expanded is used[3, 6]. Data from 74 subjects at 100Hz for sampling rate were downloaded and analyzed using Python Programming Language [11], MNE Library for EEG analysis [4] and Scikit-Learn [8] as Machine Learning Library for classification stage. For each subject there are two night of sleep recording, for a 148 total night of sleep recorded. A notch filter to clean data from the noise D.C. current (50Hz) is applied, and after, is applied another band pass filter from 1 to 50 Hz. Each data contains the labeling manually performed by the original authors of the dataset, and the labels are:

- 'Movement time': movement of subject
- 'Sleep stage 1'
- 'Sleep stage 2'
- 'Sleep stage 3'
- 'Sleep stage 4'
- 'Sleep stage ?'
- 'Sleep stage R' : Rem Stage
- 'Sleep stage W': Awake

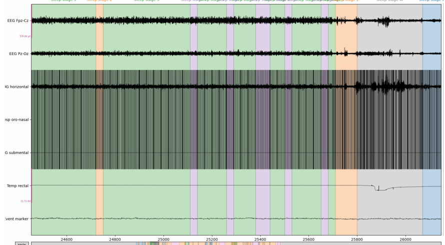

Fig. 1. Sleep Stage

The EEG instrument used for the recordings were composed of four EEG channels with bipolar references.

- Fpz-Cz
- Pz-Oz

and others additional channels for other physiological processes:

- EOG channel
- Resp oro-nasal
- EMG submental
- Temp rectal
- Event marker

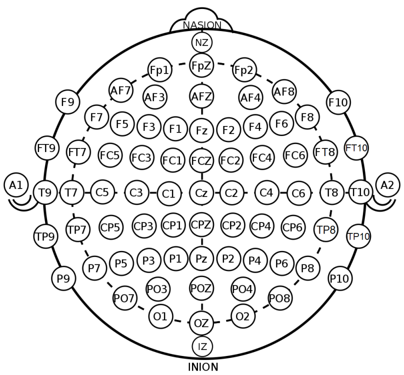

Fig. 2. 10-20 system electrodes position

This classification model is based only on four EEG channels, so other channels were excluded.

The length of each epoch to be classified was chosen at 30 seconds based on the scientific literature and manuals [9]. The signal analyzed was cut from the first drowziness moment to the wake-up moment, exluding the rest of the day. Then a list is created containing the label, the start and the end of each epoch in term of index of timepoint of the signal. Then, with Welch's method [12], for each band of the signal, the relative band power is extracted:

$$RelativeBandPower = BandPower/TotalPower \quad (1)$$

With this computation, we have the percentage of power for each band to be independent of the amplitude of the signal that can be not stable across subjects.

The bands used are defined as:

- Delta: from 1 Hz to 4 Hz
- Theta: from 4 Hz to 8 Hz
- Alpha: from 8 Hz to 12 Hz
- Beta: from 12 Hz to 30 Hz
- Gamma: from 30 Hz

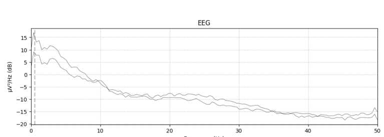

Fig. 3. Power Spectral Density

This is computed for each pair of electrodes. Finally, a dataframe containing information about the subject, recording, relative band power for each pair of eletrodes, and label was built. A total dataframe of 153953 samples were extracted. The epoch with a label not related to sleep stage were excluded:

- 'Movement time': *movement of subject*
- 'Sleep stage ?': *ambiguous epoch*
- 'Sleep stage W': awake

Finally, the dataset was composed by 106902 samples. The ratio from train and test set was fixed as 0.33 and the validation was the 0.1 portion of the test set.

Considering only the random extraction, the probability for each class are the following, normalized on the number of sample for each class.

- Sleep stage 1 0.433
- Sleep stage 2 0.207
- Sleep stage 3 0.053
- Sleep stage 4 0.123
- Sleep stage REM 0.184
- casual probability unweighted: 0.2

Algorithm 1 Pipeline

for `<raw>` **do**
  apply filters
  exclude useless channels
  find first and last events
  create epochs
  for `<epoch>` **do**
    for `<channel>` **do**
      extract power
      compute relative band power
      append to dataframe
      classify

IV. Results

In this work, four models for signal classification were used. In Fig. 5 is possible to understand the architecture of the network.

A. Multi-Layer Perceptron

First, the multilayer perceptron was built with these hyperparameters:

- 4 hidden layer: 10,30,30,8
- Activation function: 'relu'
- Solver: 'adam'
- max iter: 3000
- shuffle: True
- momentum: 0.9
- learning rate: 'adaptive'

The results for the MLP classifier were the following:

- Accuracy: 0.671
- Balanced Accuracy Score: 0.494
- F1 score: 0.659
- Recall score: 0.671
- ROC AUC score One VS Rest: 0.925
- ROC AUC score One VS One: 0.913

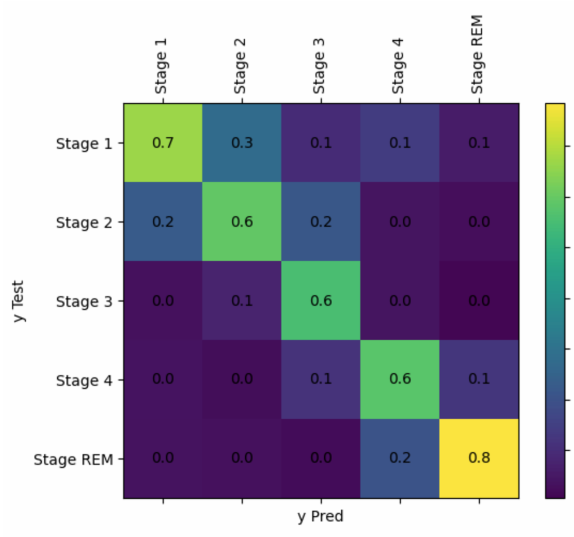

Fig. 4. Confusion Matrix

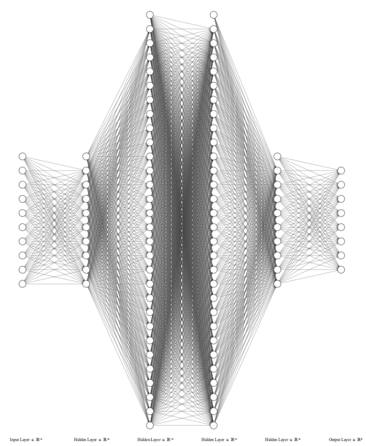

Fig. 5. Example of Network Architecture

B. Support Vector Machine

Second test was performed on a Support Vector Classifier with degree of 3 and polinomial kernel. The accuracy score was around 0.589.

C. K-Nearest Neighbors

The third test was performed using a K-Nearest Neighbors algorithm with number of neighbors = 10 and weights = 'distance'. The accuracy score was around 0.627.

D. Linear Discriminant Analysis

The fourth test was performed using Linear Discriminant Analysis with Singular Value Decomposition and the accuracy score 0.548.

E. Logistic Regression

The last test was performed using simple multi-class logistic regression and the results was 0.548.

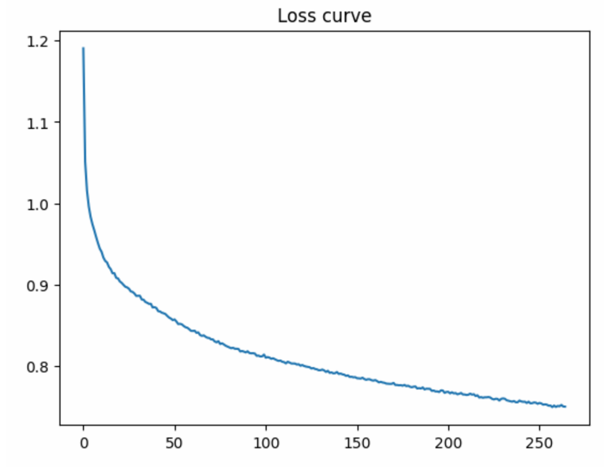

Fig. 6. Loss curve

V. Grid Search

In order to have a better classification, it was performed a Grid Search to find the best hyperparameters for multi-layer perceptron with Cross Validation = 2. The best model with this cross validation was the MLP with following hyperparameters:

- alpha=0.005
- hidden layer sizes=(8, 16, 32, 64, 8)
- learning rate init=0.0001
- max iter=300
- random state=42
- solver='sgd')

The metrics of the best model were the following:

- accuracy: 0.610
- balanced accuracy score: 0.486
- f1 score: 0.567
- recall score: 0.610
- roc auc score: 0.853
- roc aucsscore: 0.852

References

[1] Khald Ali I Aboalayon et al. "Sleep stage classification using EEG signal analysis: a comprehensive survey and new investigation". In: *Entropy* 18.9 (2016), p. 272.

[2] Eugene Aserinsky and Nathaniel Kleitman. "Two types of ocular motility occurring in sleep". In: *Journal of applied physiology* 8.1 (1955), pp. 1–10.

[3] Ary L Goldberger et al. "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals". In: *circulation* 101.23 (2000), e215–e220.

[4] Alexandre Gramfort et al. "MEG and EEG data analysis with MNE-Python". In: *Frontiers in neuroscience* (2013), p. 267.

[5] Eric R Kandel et al. *Principles of neural science*. Vol. 4. McGraw-hill New York, 2000.

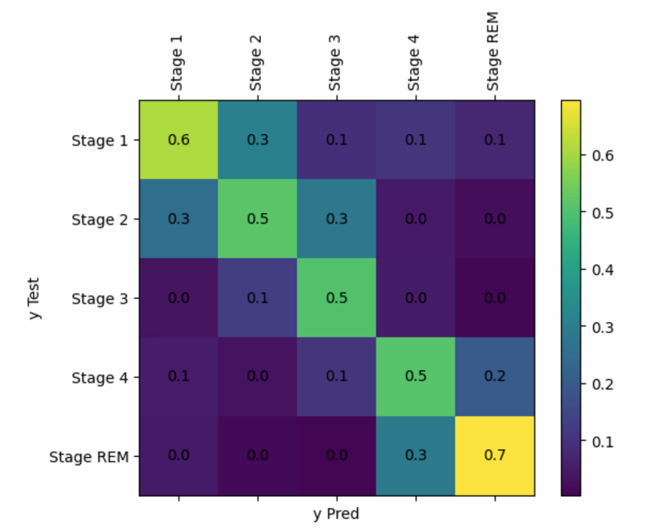

Fig. 7. Confusion Matrix with best model

[6] Bob Kemp et al. "Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG". In: *IEEE Transactions on Biomedical Engineering* 47.9 (2000), pp. 1185–1194.

[7] Hui Wen Loh et al. "Automated detection of sleep stages using deep learning techniques: A systematic review of the last decade (2010–2020)". In: *Applied Sciences* 10.24 (2020), p. 8963.

[8] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[9] A Rechtschaffen and A Kales. "Proposed supplements and amendments to'A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects', standard". In: *Psychiatry and Clinical Neurosciences* 55 (1968), pp. 305–310.

[10] Robert Stickgold and Matthew P Walker. *The neuroscience of sleep*. Academic Press, 2010.

[11] Guido Van Rossum, Fred L Drake, et al. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[12] Peter Welch. "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms". In: *IEEE Transactions on audio and electroacoustics* 15.2 (1967), pp. 70–73.