Characterization of EEG signals according to brain regions using machine learning techniques

Lai Khang Duy - USTHBI7-051

Supervisors: Sebastian Basterrech

October 4, 2019

ICT Department - University of Science and Technology of Hanoi

Table of Contents

Table of Contents

Brain

Figure 1: Human brain

Introduction

- Every vertebrates and most of invertebrates have brain.
- When it works, it generates electrical signal.
- Electroencephalography (EEG) is a method to record the neuraloscillations (often known as "brain waves") of the brain.
- Different region of the brain generates specific type of different signal.

Problem and Objective

- **Objective:** Characterize brain regions over labeled EEG signal.

Problem and Objective

- **Objective:** Characterize brain regions over labeled EEG signal.
- **Why we have to study this?**

Problem and Objective

- **Objective:** Characterize brain regions over labeled EEG signal.
- **Why we have to study this?**
    - When some area of the brain is damaged, the brain plasticity provokes other parts. It can be activated and have new functionalities in order of mitigating the impact of brain damages.

Problem and Objective

- **Objective:** Characterize brain regions over labeled EEG signal.
- **Why we have to study this?**
  - When some area of the brain is damaged, the brain plasticity provokes other parts. It can be activated and have new functionalities in order of mitigating the impact of brain damages.
  - So here is why:
    1. Analyzing the plasticity of the brain.
    2. How damaged brain provokes changes in their functionality.

Problem and Objective

- **Objective:** Characterize brain regions over labeled EEG signal.
- **Why we have to study this?**
  - When some area of the brain is damaged, the brain plasticity provokes other parts. It can be activated and have new functionalities in order of mitigating the impact of brain damages.
  - So here is why:
    1. Analyzing the plasticity of the brain.
    2. How damaged brain provokes changes in their functionality.
- Our goal is to have a general comparison between 2 well-known machine learning techniques LSTM and SVM in classifying brain regions.

Table of Contents

Brain electrical signals

More specific

- Brain is a complex organ, contain over 86 bilions neurons.
- Each neuron connects to thousands of others.
- Communicate over synapses electrical signal.

Electroencephalography

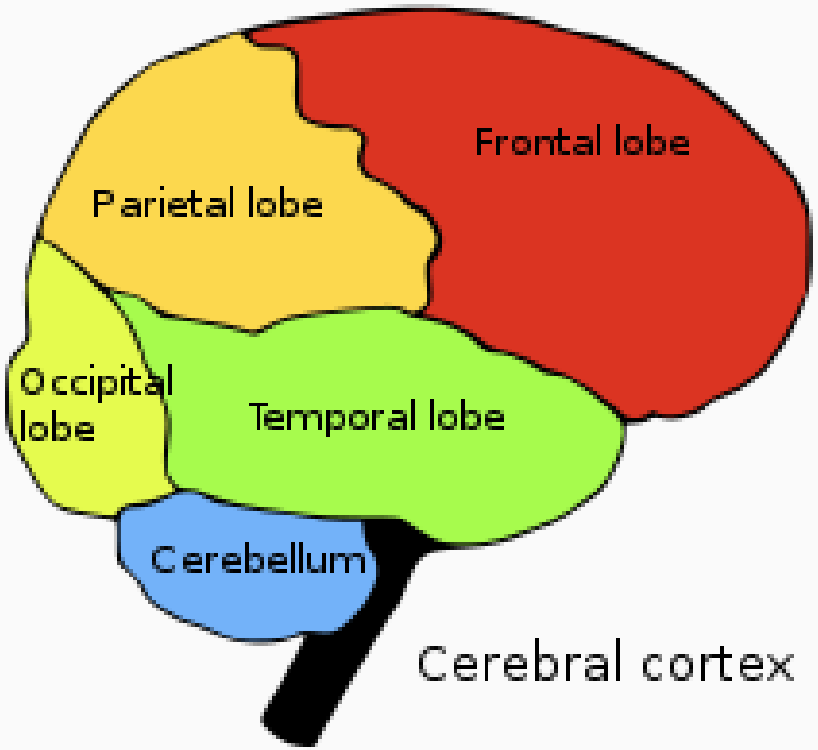

Figure 2: Region of the brain

- Frontal lobe (F) (Red)

Electroencephalography

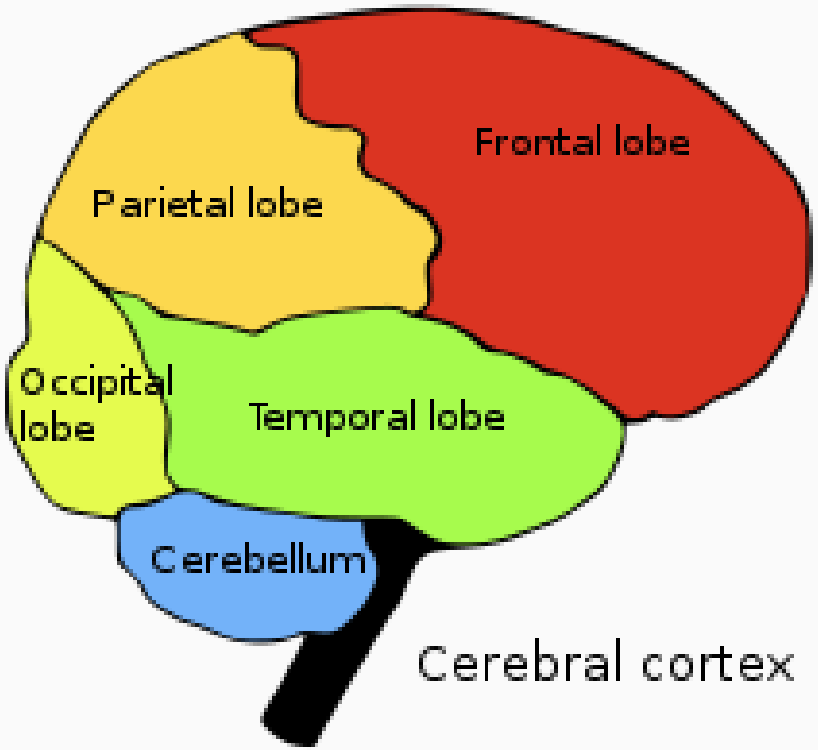

Figure 2: Region of the brain

- Frontal lobe (F) (Red)
- Cerebellum (C) (Cyan)

Electroencephalography

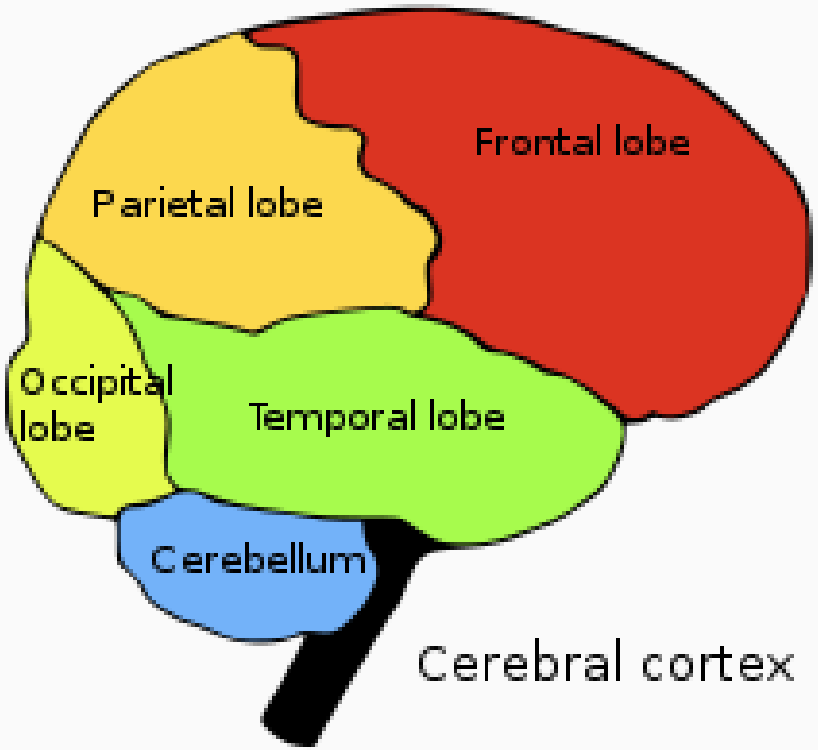

Figure 2: Region of the brain

- Frontal lobe (F) (Red)
- Cerebellum (C) (Cyan)
- Parietal lobe (P) (Yellow)

Electroencephalography

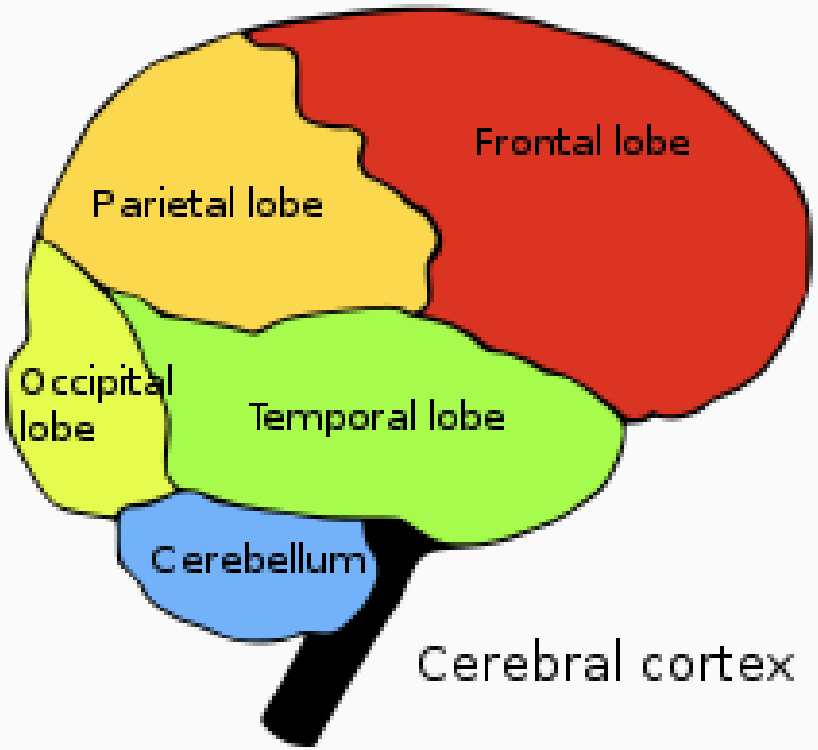

Figure 2: Region of the brain

- Frontal lobe (F) (Red)
- Cerebellum (C) (Cyan)
- Parietal lobe (P) (Yellow)
- Occipital lobe (O) (Emerald)

Electroencephalography

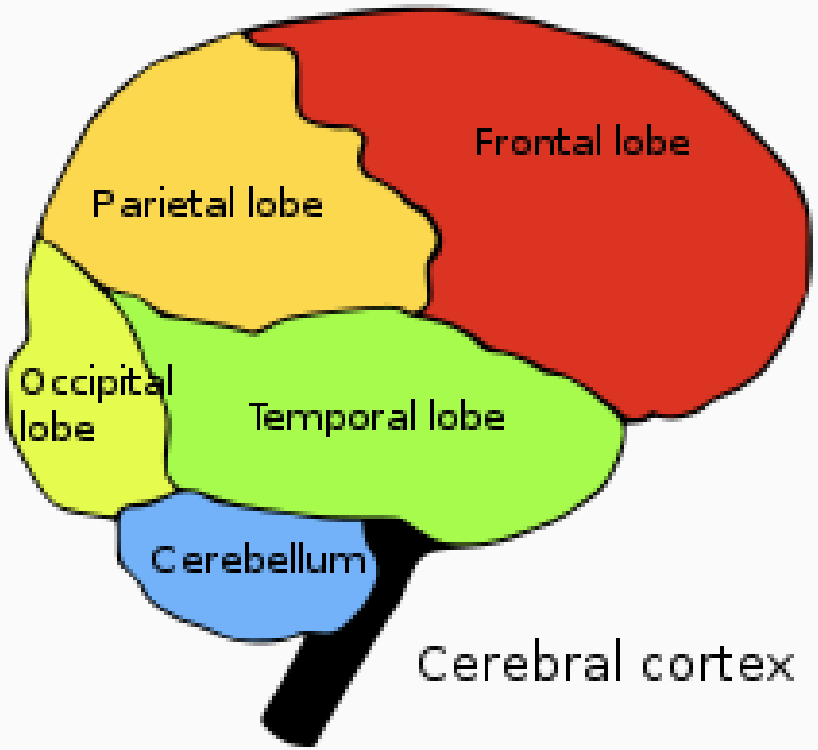

Figure 2: Region of the brain

- Frontal lobe (F) (Red)
- Cerebellum (C) (Cyan)
- Parietal lobe (P) (Yellow)
- Occipital lobe (O) (Emerald)
- Temporal lobe (T) (Green)

Electroencephalography

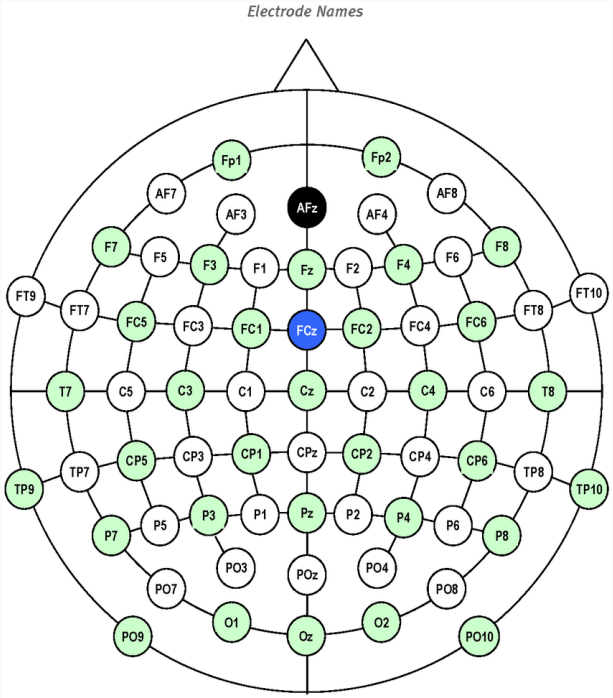

Figure 3: Full electrode map system

Electroencephalography

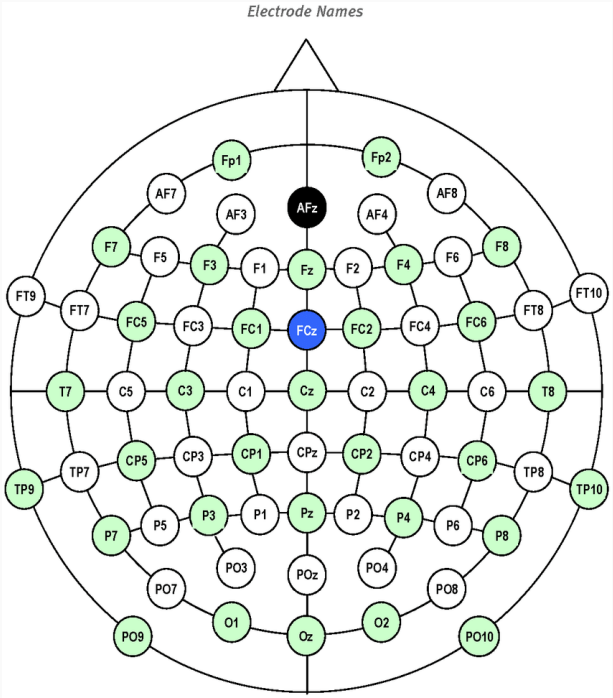

Figure 3: Full electrode map system

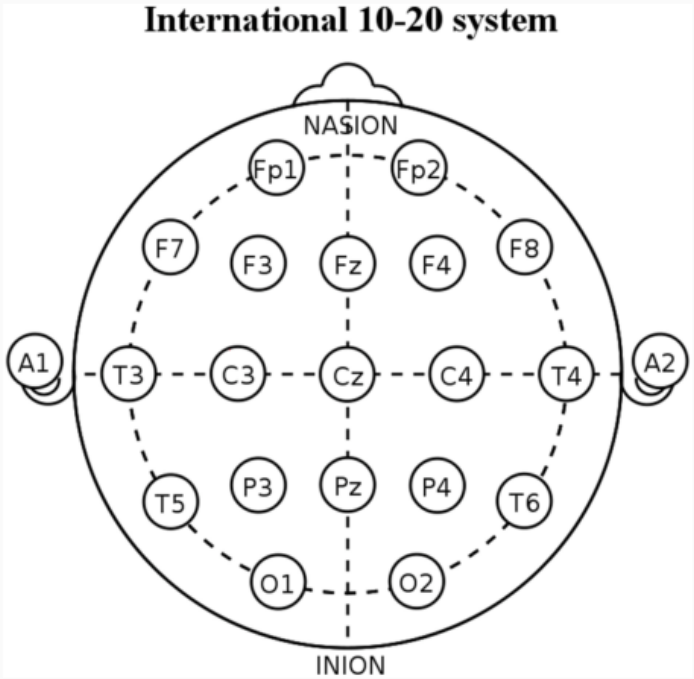

Figure 4: 19 electrodes map system

Electroencephalography

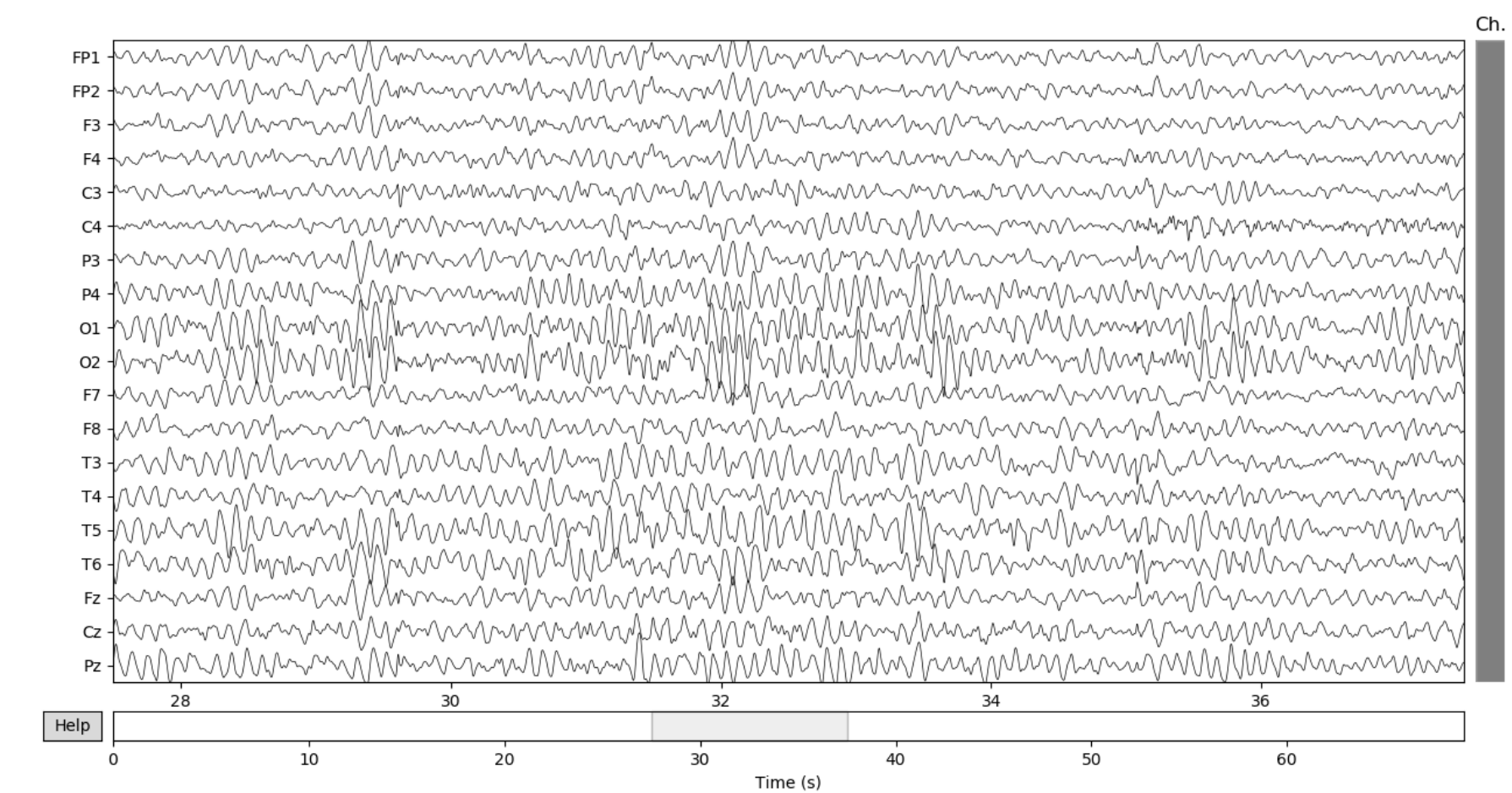

Figure 5: 19 EEG signal example

Table of Contents

EEG signals pre-processing

To have the data ready for the machine learning part, we have to get through:

- Noise reduction
- Learning set creation

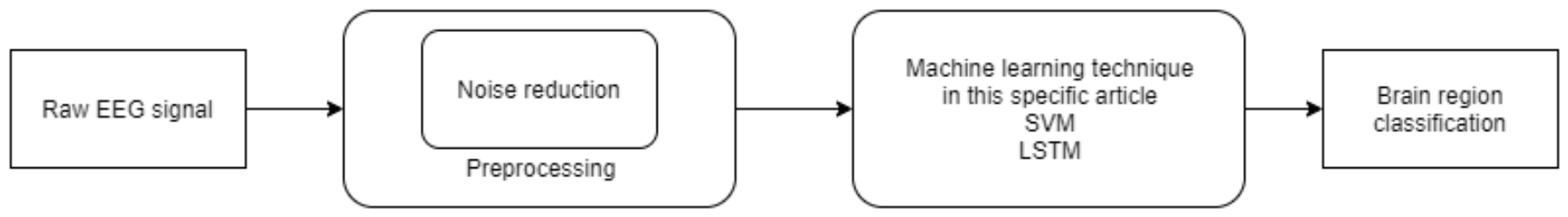

Figure 6: Project sequences breakdown

Noise reduction

The technique that we use to denoise in this study is Z-score.

$$z = \frac{x - \mu}{\sigma}$$

Noise reduction

The technique that we use to denoise in this study is Z-score.

$$z = \frac{x - \mu}{\sigma}$$

Why ?

Noise reduction

The technique that we use to denoise in this study is Z-score.

$$z = \frac{x - \mu}{\sigma}$$

Why ? Because EEG signal very stable. Any significant changes is considered as outliers. In this case, I want to consider around 5% of the data we have is outliers.

Noise reduction

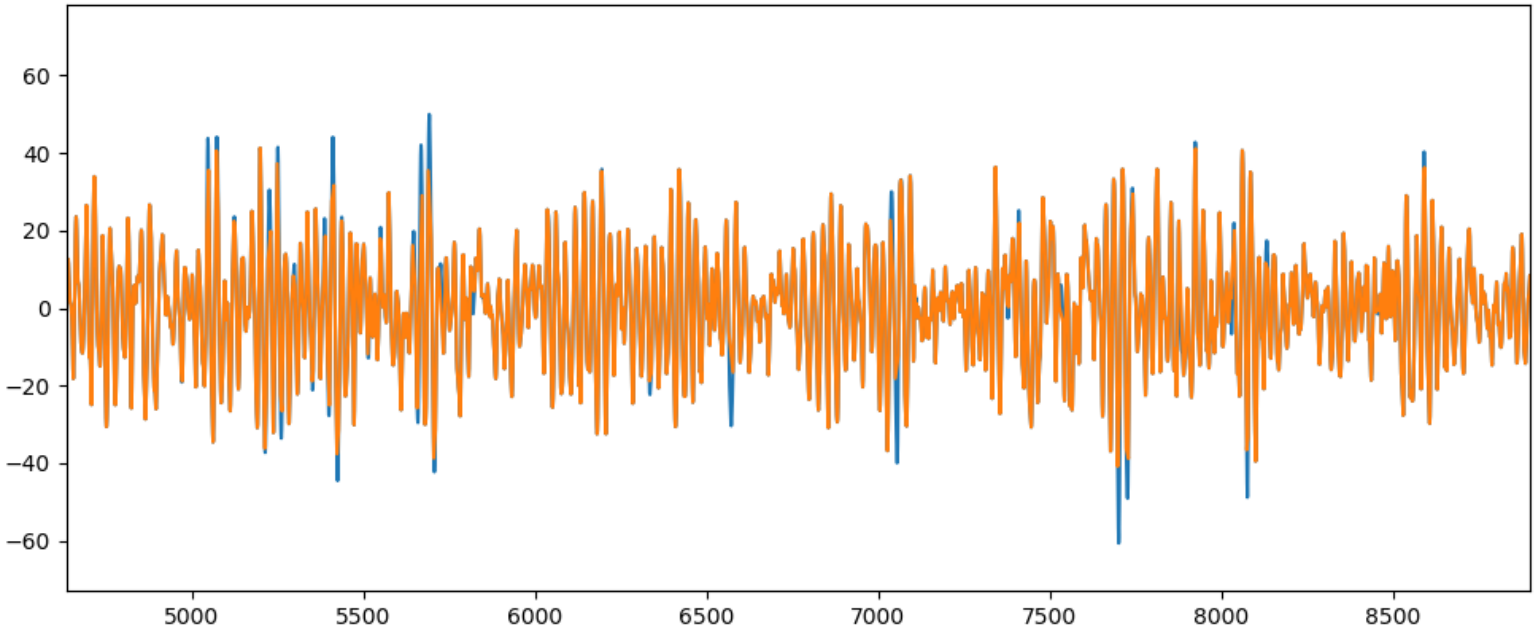

Figure 7: FP1 before and after apply Z-score

Noise reduction

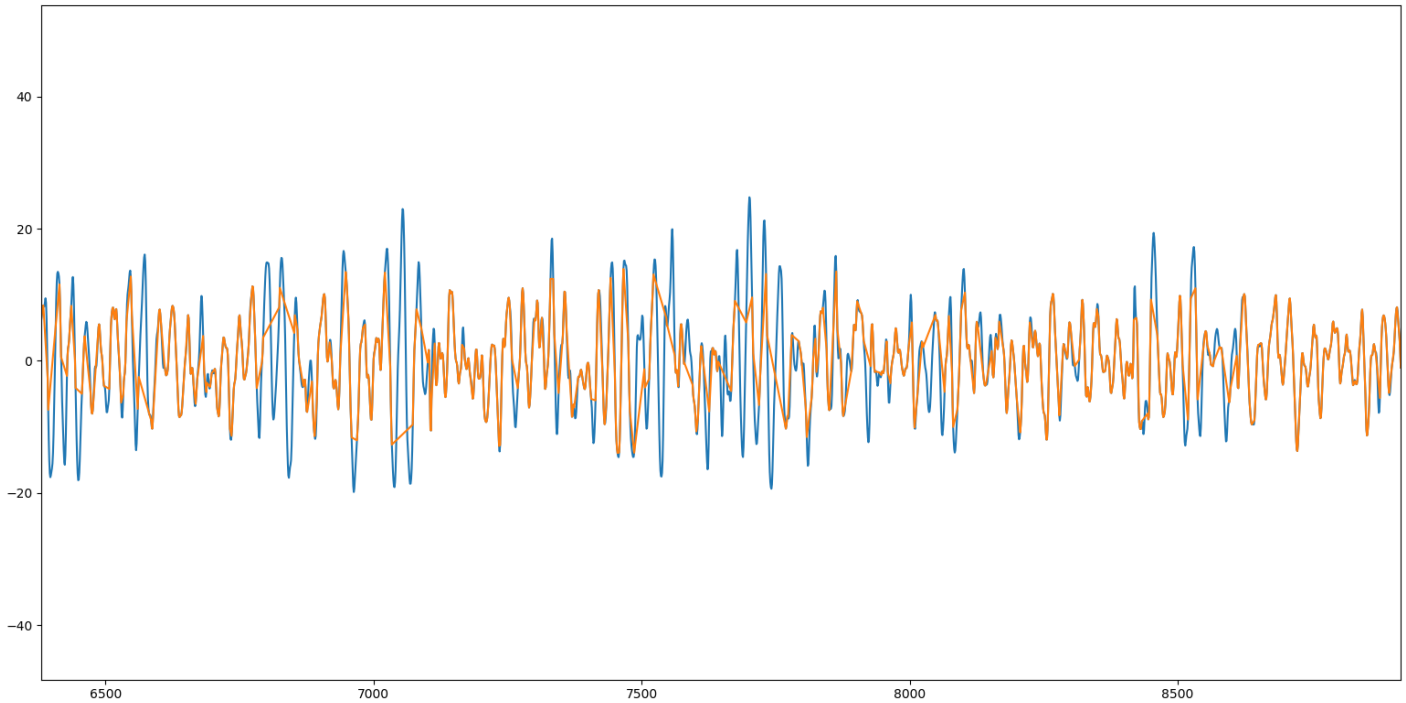

Figure 8: Closer look at FP1 before and after apply Z-score

Learning set creation

- 19 electrodes

Learning set creation

- 19 electrodes
- 4 classes:
  1. Class 0: including FP1-AVE, FP2-AVE, F3-AVE, F4-AVE, F7-AVE, F8-AVE, Fz-AVE
  2. Class 1: including C3-AVE, C4-AVE, Cz-AVE
  3. Class 2: including T3-AVE, T4-AVE, T5-AVE, T6-AVE
  4. Class 3: including P3-AVE, P4-AVE, Pz-AVE, O1-AVE, O2-AVE

Learning set creation

- 19 electrodes
- 4 classes:
  1. Class 0: including FP1-AVE, FP2-AVE, F3-AVE, F4-AVE, F7-AVE, F8-AVE, Fz-AVE
  2. Class 1: including C3-AVE, C4-AVE, Cz-AVE
  3. Class 2: including T3-AVE, T4-AVE, T5-AVE, T6-AVE
  4. Class 3: including P3-AVE, P4-AVE, Pz-AVE, O1-AVE, O2-AVE
- 2 datasets with the same signal:
  1. W1: 200 time steps
  2. W2: 400 time steps

Learning set creation

- 19 electrodes
- 4 classes:
  1. Class 0: including FP1-AVE, FP2-AVE, F3-AVE, F4-AVE, F7-AVE, F8-AVE, Fz-AVE
  2. Class 1: including C3-AVE, C4-AVE, Cz-AVE
  3. Class 2: including T3-AVE, T4-AVE, T5-AVE, T6-AVE
  4. Class 3: including P3-AVE, P4-AVE, Pz-AVE, O1-AVE, O2-AVE
- 2 datasets with the same signal:
  1. W1: 200 time steps
  2. W2: 400 time steps
- Shuffle whole dataset after divide classes and time windows.

Learning set creation

2 same datasets with different time window

- W1: 200 time steps each datapoint
- W2: 400 time steps each datapoint

Table of Contents

LSTM-RNN

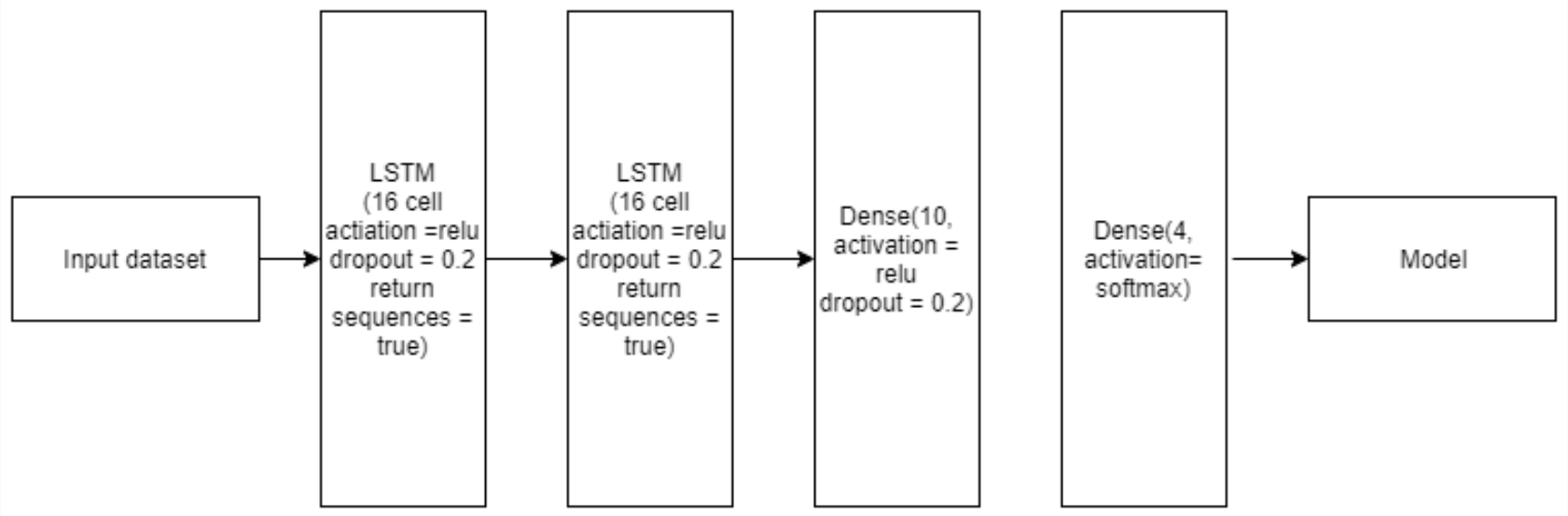

Figure 9: LSTM neural network breakdown

SVM

- Implement SVC in scikit-learn
- Parameter:
    1. Gamma: scale
    2. C: 100

Table of Contents

LSTM-RNN

	W2 = 400
Epoch 1/2	acc = 0.3173, val_acc = 0.4255
Epoch 2/2	acc = 0.3387, val_acc = 0.4326

Table 1: Accuracy using LSTM with window W2=400 time steps

	W1 =200
Epoch 1/2	acc = 0.1941, val_acc = 0.2626
Epoch 2/2	acc = 0.2662, val_acc = 0.3805

Table 2: Accuracy using LSTM with window W1=200 time steps

SVM

W2 = 400	W1 =200
accuracy = 0.72	accuracy = 0.575

Table 3: Accuracy using SVM with windows W1=200 and W2=400 time steps

Table of Contents

Discussion

- SVM has better accuracy than LSTM-RNN.
- window W2 has better accuracy than window W1.
- According to our experiments, the highest accuracy was 72%.

Table of Contents

Conclusion

The result of this experiment still not as we expected.

We have figured out some of the reasons might affect the result:

- Lack of data.

Conclusion

The result of this experiment still not as we expected.

We have figured out some of the reasons might affect the result:

- Lack of data.
- ECG and EOG still have not be removed.

Conclusion

The result of this experiment still not as we expected.

We have figured out some of the reasons might affect the result:

- Lack of data.
- ECG and EOG still have not be removed.
- The filter is still not perfect.

Future works

It is still possible to improve the classifier.

- Collect more data.
- Improve the noise filter.