Assessing Neurological States from Physiological Signals

Sama Amr Gouda
900211296

Mona Ibrahim
900212749

Shaymaa Nabil
900214260

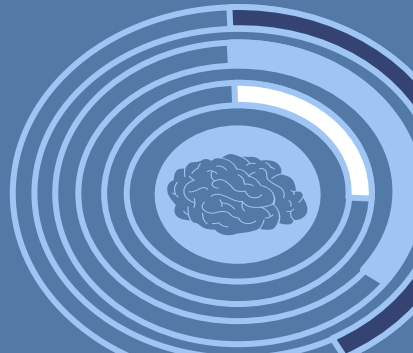
Table of Contents

01 Introduction

Introduction

Traditionally, an EEG (electroencephalography) is used to measure the electrical activity in the brain, helping in diagnosing the brain condition.However, it's resource intensive and requires specialized equipment

Non-EEG physiological signals examination has been introduced where it "evaluates brain and nervous system functioning"



Problem Statement

Classify these collected physiological signals into either one of the four neurological states:

- **Physical stress**
- **Cognitive stress**
- **Emotional stress**
- **Relaxation**

02 Dataset Description

Dataset Description

- Retrieved from Physionet and is collected at the Quality of Life Laboratory at University of Texas at Dallas
- Collected using non-invasive wrist worn biosensors
- Measure 5 different attributes
- 20 subjects, where each had to undergo a series of 7 stages

Dataset Description

Electrodermal activity (EDA)

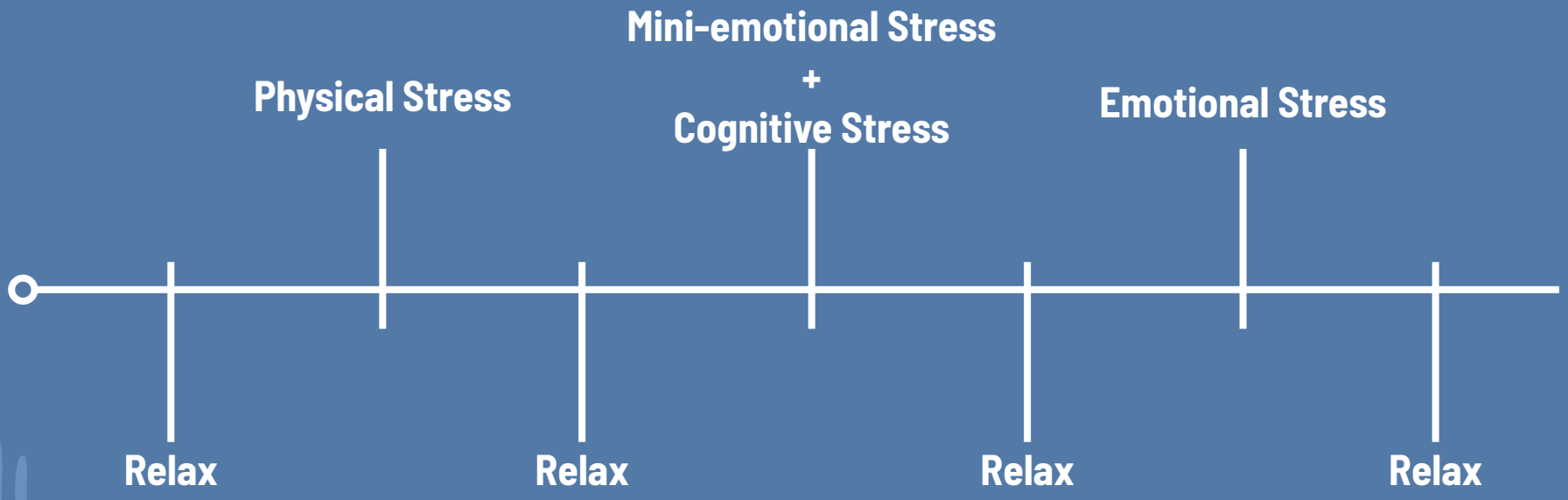
Temperature

Acceleration

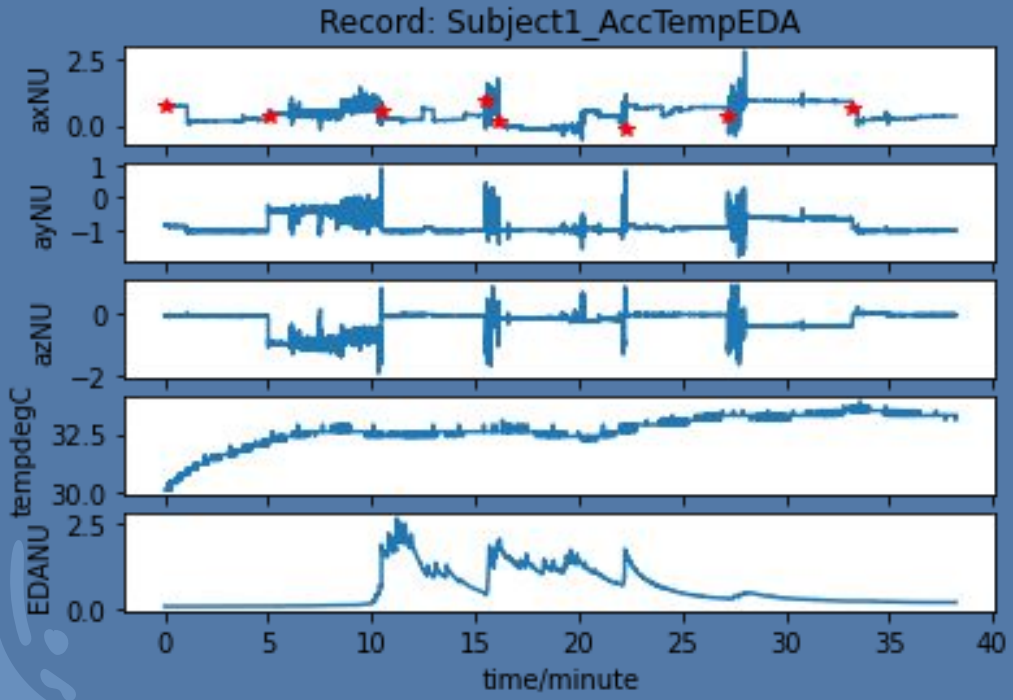
Heart rate (HR)

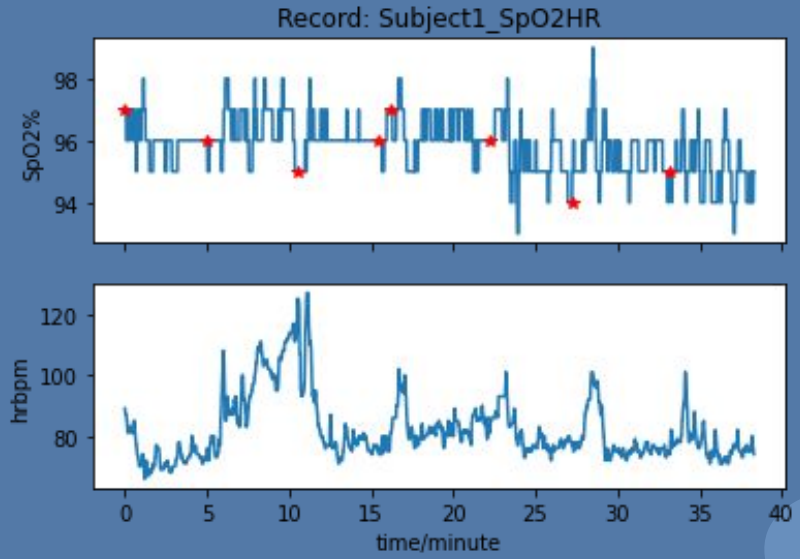
Arterial oxygen level (SpO2)

Dataset Description



Dataset Description





03 Dataset Preprocessing

Data preprocessing

01

Signal Processing I

02

Signal resizing

03

Feature Extraction

01

Initial Signal Processing

Signal Types and Sampling Rates

- **AccTempEDA File:**
  - **Sampling Rate:** 8 Hz (8 samples/second).
- **SpO2HR File:**
  - **Sampling Rate:** 1 Hz (1 sample/second).

Annotation Parsing

- **Sample Indices:** Define start and end points for each stage (8 Hz).
- **Stage Names:** Labels such as 'Relax', 'PhysicalStress'.

01

Initial Signal Processing

Segmentation Process

- **For each stage (except last):** Extract data from sample[j] to sample[j+1].
- **For the last stage:** Extract from sample[j] to the end of the recording.

Result:

- Signals are grouped into dictionaries by stage (e.g., Relax, PhysicalStress).

02

Signal Resizing

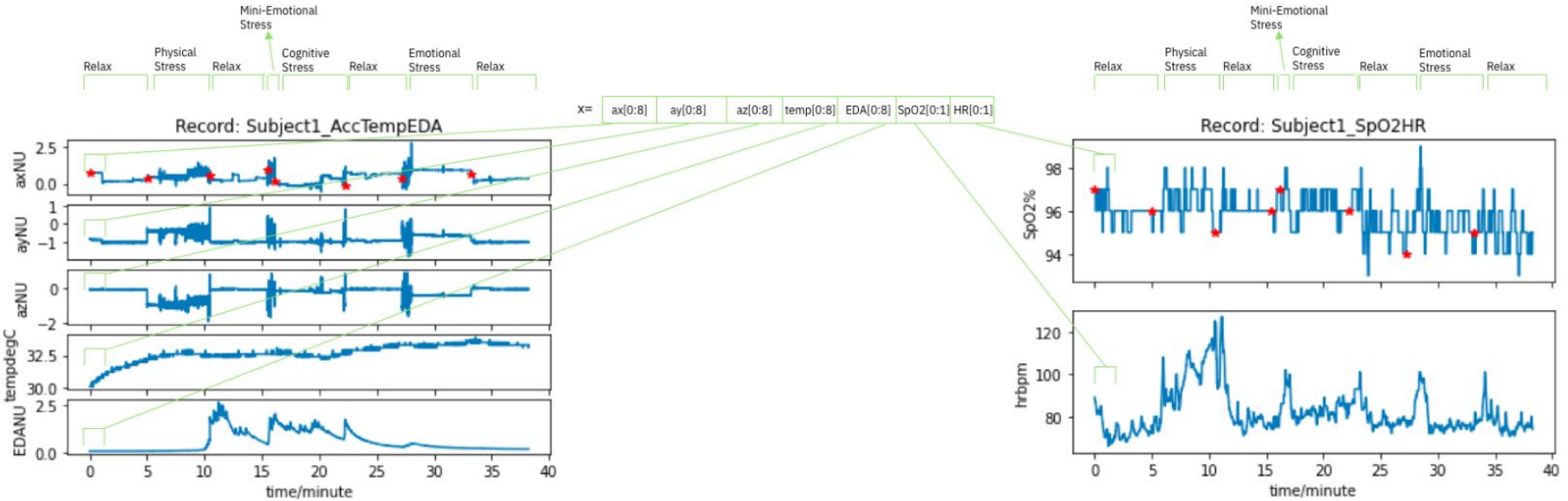
- **Objective:** Ensure consistent lengths across subjects and stages.
- Resized based on minimum lengths for each stage:
  - **AccTempEDA Signals:** e.g., 2384 samples for 'Relax'.
  - **SpO2HR Signals:** e.g., 298 samples for most stages.
- **Approach:** Truncate longer signals; pad shorter ones.

03

Feature Extraction

- Combine data into a single vector per observation:
  - **AccTempEDA Signals:** 8 samples × 5 signals = 40 values.
  - **SpO2HR Signals:** 1 sample each from SpO2 and HR = 2 values.
- **Total Features per Observation:** 40 + 2 = **42 Features**.

Feature Vector



04 Modeling and Results

Approach for modeling

Goal: Classify Activity Stages
Stages: Relax, Physical Stress, Emotional Stress, Cognitive Stress.

Methodology:

1. Train machine learning models for each subject individually.
2. Use labeled feature vectors for training.
3. Evaluate performance for accurate classification.

Model Results

99.18%

Random Forest

Using Cross Validation without PCA

	Model	Cross-Validation	Leave-one out	PCA	Accuracy
0	Multi-Layer Perceptron	Yes	---	Yes	98.85
1	Multi-Layer Perceptron	--	Yes	Yes	66.42
2	Multi-Layer Perceptron	Yes	--	NO	66.42
3	KNN	Yes	---	Yes	93.68
4	KNN	---	Yes	Yes	71.62
5	KNN	Yes	--	NO	91.48
6	Random Forest	Yes	---	Yes	98.80
7	Random Forest	---	Yes	Yes	77.55
8	Random Forest	Yes	---	NO	99.18
9	Decision Tree	Yes	---	Yes	93.93
10	Decision Tree	---	Yes	Yes	77.19
11	Decision Tree	Yes	---	NO	95.64
12	Naive Bayes	Yes	---	Yes	90.51
13	Naive Bayes	---	Yes	Yes	72.53
14	Naive Bayes	Yes	---	No	91.76
15	FLDA	Yes	---	Yes	91.86
16	FLDA	---	Yes	Yes	77.24
17	FLDA	Yes	---	No	91.89
18	LS	Yes	---	Yes	64.36
19	LS	---	Yes	Yes	41.71
20	LS	Yes	---	No	64.28

Due to robustness against overfitting and capturing nonlinear interactions within the data.

Feature Importance

- Exploring the feature importance for Random Forests.
- Used Gini importance
- The most important feature in the feature vector feature 17 that is related to “az” which is part of the Electrodermal activity

```
Max importance feature:  17
Importance:  0.0955849550981549
Value Related to az
```

Feature	Importance
1	0.003918
2	0.008186
3	0.004017
4	0.005090
5	0.005818
6	0.004335
7	0.005105
8	0.001730
9	0.003299
10	0.001519
11	0.005849
12	0.002760
13	0.004015
14	0.001629
15	0.003640
16	0.003829
17	0.095585
18	0.088780
19	0.057243
20	0.046901
21	0.061616
22	0.057379
23	0.070388
24	0.048101
25	0.027703
26	0.037728
27	0.019445

05

Graphical User Interface (GUI)

GUI

Input = Subject Number

+

Randomly select signal vector at time ‘t’ that contains a combination of the different attributes for that particular subject

Output = classification of the state this vector is at

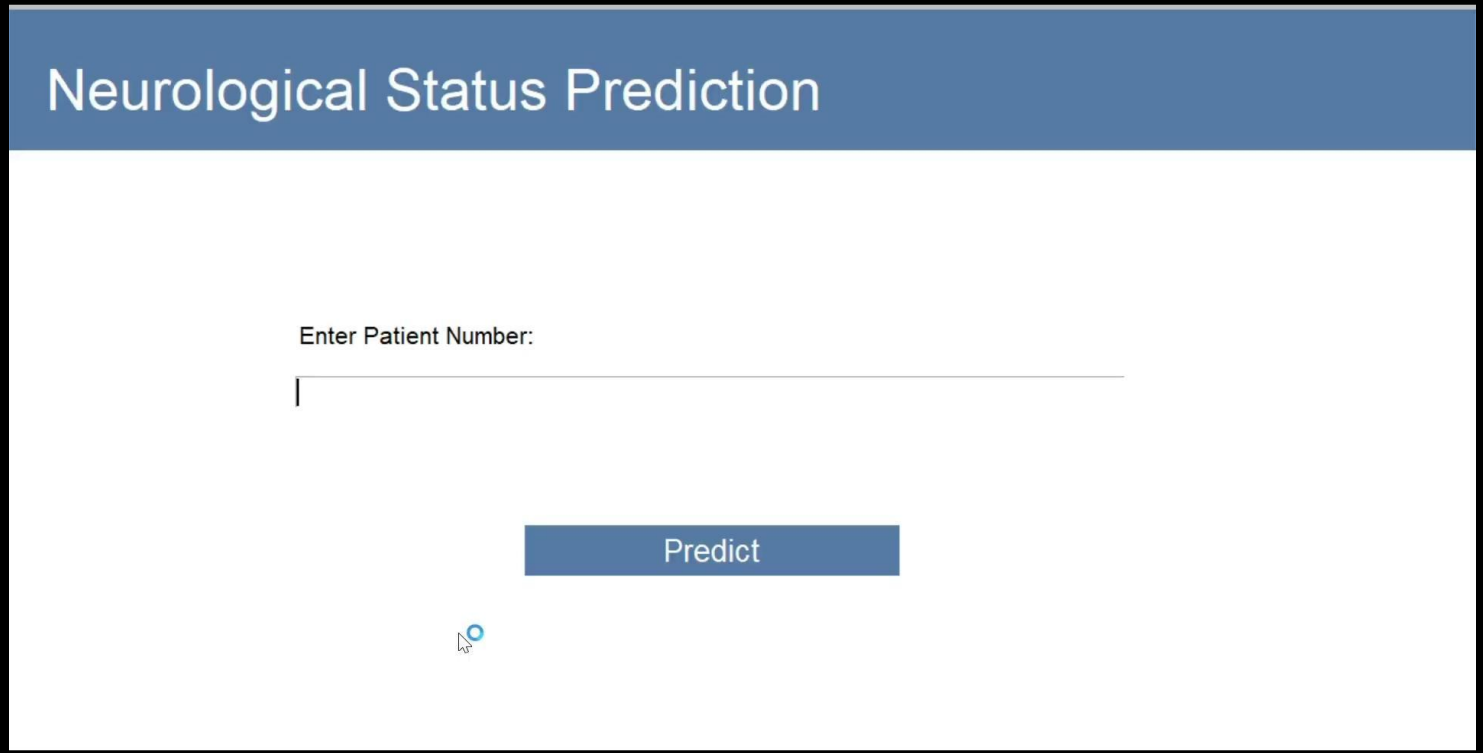
+

comparing it to the ground truth

GUI

- Tkinter library was used to compile this GUI
- Better help doctors and Lab-technicians in identifying signals faster, minimizing the inference time where the results are just a click away; in addition to, minimizing the human error
- For future work, we would like to upgrade such an application to take a subject's file as an input even if it isn't contained within the dataset, so that it's more dynamic and can be generalized and used over more numbers of patients.

GUI


Neurological Status Prediction
Enter Patient Number:
Predict

06 Conclusion

Random Forest achieved the highest accuracy in classifying neurological states using physiological signals.

Future work could focus on

- Expanding the dataset
- Refining the classifiers
- Improving the GUI to further enhance usability and accuracy.

References

- Cleveland Clinic Medical. (2022, April 1). Neurological exam. Cleveland Clinic. https://my.clevelandclinic.org/health/diagnostics/22664-neurological-exam

- EEG (electroencephalogram). (2024, May 29). Mayo Clinic. https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875