Exercise 2: Physiological signals in affective computing

In the first exercise, you extracted information from a series of sensors about a person's physical state. In this exercise, you will use sensor values to learn more about a person's internal state. For this, you will need to apply concepts that you have encountered in this course related to biomedical signals and their application within the field of affective computing.

Specifically, you are provided with electrocardiograms (ECG), electroencephalograms (EEG), electrodermal activities (EDA) (also known as galvanic skin response (GSR)) and facial landmark trajectories (EMO) of 44 participants as they watched different affective movie clips from the ASCERTAIN dataset. Your goal is to build a system that recognizes a participant's current emotional state as described by valence and arousal.

You will train a random forest (RF) classifier with a set of suitable features, which you can extract from the given data, and the corresponding valence and arousal levels that participants have assigned themselves after watching each clip. Together with your partner, you will create a short report that summarizes and interprets your findings as the final deliverable.

1 Data

Note: **All data provided is confidential**. You are not allowed to share or use any parts of this data beyond the scope of this exercise. The data have not been released publically and you are not allowed to distribute them.

1.1 Access

The data ("Exercise2_data.zip") can be downloaded from the "Announcements" *Microsoft Teams Ubiquitous Computing Spring 2021* channel.

After extracting the data, you will find a folder with 44 Python Pickle files. Each file stores the data for selected recordings of a single participant. We already discarded any samples with known issues during the recording process.

The data for a participant can be loaded as follows:

```python
import pickle
# "import pickle5 as pickle" in case you are working with Python < 3.8

# Load data
with open("participant_01.pkl","rb") as f:
    participant_data = pickle.load(f)
```

1.2 Data Composition

Each pickle file stores a dictionary containing the data for a single participant.
The dictionary has the following entries:

- "ID": ID of participant.
- "FS_ECG" Sampling frequency for ECG data in Hz.
- "FS_EEG" Sampling frequency for EEG data in Hz.
- "FS_GSR" Sampling frequency for GSR data in Hz.
- "FS_EMO" Sampling frequency for EMO data in Hz.
- "recordings": Dictionary storing corresponding data entries for each clip ID
  - "arousal": self-rating for arousal (0–6)
  - "valence": self-rating for valence (−3–3)
  - "ECG": $L \times 3$ array where L is the length of the recording in samples; the columns contain a timestamp in milliseconds for each sample, as well as the ECG signals from the right and left hand respectively.
  - "EEG": $L \times 8$ array where L is the length of the recording in samples; the columns contain the eight channels of the EEG.
  - "GSR": $L \times 2$ array where L is the length of the recording in samples; the columns contain a timestamp as well as the GSR signal in Ohm.
  - "EEG_features": array of size 88; contains the extracted features from the recorded EEG signals. The features consist of the (1–6) statistical measurements over the channel, (7) average of the first derivative, (8) proportion of negative differential samples, (9) mean number of peaks, (10) mean derivative of the inverse channel signal, (11) average number of peaks in the inverse signal for each channel.
  - "EMO": $L \times 22$ array where L is the length of the recording in samples; the labels for the 22 columns (face landmark features) are as follows:

01: Frame number or milliseconds	12: deformation of the right lid
02: vertical deformation of the upper lip	13: deformation of the left lid
03: vertical deformation of the lower lip	14: neutral state assuming a neutral frontal initial frame
04: horizontal deformation of the left lip corner	15: happy state assuming a neutral frontal initial frame
05: vertical deformation of left lip corner	16: surprised state assuming a neutral frontal initial frame
06: horizontal deformation of the right lip corner	17: angry state assuming a neutral frontal initial frame
07: vertical deformation of the right lip corner	18: disgusted state assuming a neutral frontal initial frame
08: deformation of the right eyebrow	19: fearful state assuming a neutral frontal initial frame
09: deformation of the left eyebrow	20: sad state assuming a neutral frontal initial frame
10: deformation of the right cheek	21: x dimension of the head pose
11: deformation of the left cheek	22: y dimension of the head pose

2 Tasks

For the following tasks, you should report your results and plot figures. Numerical results are often best highlighted in the form of tables. Describe your findings and include brief interpretations of the results for each task in the report.

2.1 Feature extraction and data analysis (10 points total)

For this task, your goal is to pre-process the data and extract discriminative features for the classification of valence and arousal. Following the paper you read in *Reading Assignment 3*, whenever we mention **statistical measurements**, we refer to 1) mean (μ), 2) standard deviation (std σ), 3) skewness, 4) kurtosis, as well as the percentage of times the value is 5) above $\mu + \sigma$ and 6) below $\mu - \sigma$ of the corresponding feature or signal.

2.1.1 ECG (4.5 points)

- Plot the Power Spectral Density (PSD) for the Lead I ECG signal of the first participant watching Clip 1 from 0–40 Hz. (0.5 points)
- Propose and reason about, implement and apply a suitable low-pass filter to remove noise. Plot the filter's frequency response. Plot the filtered ECG signal in the time-domain. (1 point)
- Using the algorithm for artifact detection introduced in Lecture 6 or another suitable approach, compute the percentage of ECG data that is flagged as having artifacts. Select one such ECG signal that contains artifacts and plot it. (1 point)
- For the last 50 seconds of each recording, extract the following features as mentioned in the paper (1 point):

01: ten low frequency ([0–2.4] Hz) power spectral densities	04: **statistical measurements** over heart rate
02: four very slow response ([0–0.04] Hz) PSDs	05: **statistical measurements** over heart rate variability
03: **statistical measurements** over inter beat intervals	

- Propose and implement five additional features that you believe are helpful for emotion recognition. For each of the five, state a (very) brief justification for the feature. (1 point)

2.1.2 EMO (1 point)

- Compute the **statistical measurements** for the following signals as features:

01: vertical deformation of the upper lip	07: deformation of the right eyebrow
02: vertical deformation of the lower lip	08: deformation of the left eyebrow
03: horizontal deformation of the left lip corner	09: deformation of the right cheek
04: vertical deformation of left lip corner	10: deformation of the left cheek
05: horizontal deformation of the right lip corner	11: deformation of the right lid
06: vertical deformation of the right lip corner	12: deformation of the left lid

2.1.3 EDA (3.5 points)

- Plot the Power Spectral Density (PSD) for the EDA signal of the second participant watching Clip 1. (0.5 points)
- Propose and reason about, implement and apply a suitable low-pass filter to remove noise. Plot its frequency response. Plot the filtered EDA signal in the time domain. (1 point)
- For the last 50 seconds of each recording, compute the following features as mentioned in the paper (1 point):

01: mean skin resistance
02: mean of first derivatives of skin resistance
03: mean of absolute values of first derivatives of skin resistance
04: mean first derivative for negative values only
05: percentage of time with negative first derivative
06: standard deviation of skin resistance
07: number of local minima in the skin conductance signal
08: average rising time of the GSR signal
09: power density estimates; 4 sub-bands in the [0-0.4] Hz band
10: standard deviation of skin conductance
11: mean of first derivatives of skin conductance
12: mean of absolute values of first derivatives of skin conductance
13: mean of absolute values of second derivatives of skin conductance
14: number of local minima in the skin resistance signal
15: log power density estimates; 10 sub-bands in the [0-2.4] Hz band

- Propose and implement five additional features that you believe are helpful for emotion recognition. For each of the five, state a (very) brief justification for the feature. (1 point)

2.1.4 Valence and arousal (1 point)

- Calculate the Pearson correlation coefficient between valence and arousal. Interpret the obtained value. (0.5 points)
- Split the arousal levels into two classes: low = $\{0, 1, 2, 3\}$ and high = $\{4, 5, 6\}$. Split the valence levels into two classes: low = $\{-3, -2, -1, 0\}$ and high = $\{1, 2, 3\}$. Briefly describe the resulting class distributions. (0.5 points)

2.2 Classification (6 points)

Using all the features or a subset of the features you extracted as well as the EEG features provided in the pickle files, classify arousal and valence classes (low, high) using a Random Forest Classifier.

- Report the obtained accuracy, as well as precision, recall, and F1-score for each class for:
    1. a leave-one-clip-out-validation scheme for the classification of low and high arousal and valence samples (i.e., across video clips). (1.5 points)
    2. a leave-one-participant-out-validation scheme (i.e., across participants). (1.5 points)
- Plot the combined confusion matrix (sum of confusion matrices of each cross-validation split) for each experiment. (1 point)
- Which features would you select if you could only make use of 10? Justify your selection. (1 point)
- Implement the Zero Rule algorithm (ZeroR, which always predicts the majority class in the training set) as a baseline and compare its performance with the Random Forest Classifier. (1 point)

2.3 Discussion (4 points)

Interpret your results.

- Which of your features and the recorded sensing modalities are most relevant for emotion recognition? Are they applicable for unobtrusive and continuous tracking in real life? (1 point)
- How does your proposed classifier generalize across people and clips? (1 point)
- Does your classifier perform better for very high or low valence and arousal self-ratings? (1 point)
- Brainstorm two additional sensing modalities that would help improve classification accuracy and justify why. (1 point)

3 Submission and Grading

The results of this exercise should be submitted in the form of a report. The deadline for submission through Moodle is May 31, 2021. Please submit a single report together with your partner.

The report should be written using the two-column ("sigconf" option) ACM Primary Article Template (Overleaf, LaTeX, Word). The report should be 2–4 pages long (including tables and figures, excluding the section for references).

The report should be self-contained and include all requested results and figures. Please include both group members' nethz usernames in the filename as such: *"{username1}_{username2}_report.pdf"*.

Additionally, your submission should include the code that produces your results. For this, please create a Jupyter Notebook that contains all code. Use markup cells in the notebook to add headings, spacings, and text as appropriate. For all calculations, highlight the final results as they occur in your report. State all assumptions you make with inline comments.

The code should be self-contained and executable on another machine. This means that any software that cannot be installed via the standard package installer pip must be submitted with your code. The code should be submitted in a zipped folder named *"{username1}_{username2}_source.zip"*. Please do not submit any data.

The submission deadline is a hard deadline and we cannot accept any late submissions. This includes technical problems not caused by ETH. We encourage you to submit a first version 24 hours before the actual deadline.

3.1 Grading

The grading will be based on your submitted report and code. While we encourage you to achieve strong classification results if possible, the grading will focus less on the absolute scores achieved and more on their interpretation, their presentation, and the rationale you gave to derive them. Please allocate your time accordingly.

You will receive points depending on the quality and extent of the answer to each specific task. The awarded points will be added to the points you have collected in Exercise 1 and cumulatively contribute 50% towards your total grade for this course.

3.2 Rules

ETH rules on graded work apply. This includes but is not limited to: You are not allowed to share any aspects of your work outside of your group (this includes ideas).

You are allowed to make use of additional resources but you must explicitly identify external work in the code and in the report via proper referencing. Please respect copyrights and the basic rules concerning plagiarism.

4 Tips and Tricks

4.1 Suggested Python Libraries

While you can make use of Python libraries (including the ones mentioned below) for pre-processing, feature extraction and classification, we discourage you from applying them without understanding their underlying functionalities. For many of the requested features, you will have a better understanding and control by implementing their calculation from a lower level.

- Scikit-learn: Library for various machine learning models and tools.
- SciPy: Popular Python library for signal processing.
- Libraries for processing of physiological signals include
  - Biosppy
  - Neurokit2
  - cvxEDA

4.2 Additional resources

In case you are unsure about some of the concepts mentioned in this exercise, we have compiled a list of additional papers, tutorials, articles and blog posts you can reach out to. We highly encourage you to conduct further readings on your own.

- Random Forest Classification
- Group k-fold cross validation
- Precision, Recall, F1-score
- Welch's method

4.3 Additional remarks

- For the computation of the Power Spectral Density (PSD), you can use Welch's method. Abadi et al. [1] (Section 6.2.2) have a more detailed description about the PSD related features.
- You may optimize your hyperparameters directly for the simple cross-validation. Please be aware that this implies some overfitting on the given validation set and method and thus, your model might perform slightly worse on unseen data. To be technically correct, you would follow a nested-cross-validation strategy. You can ignore this for the sake of saving computation time for this exercise.

  More on nested cross-valiation here.

5 Any Questions?

Please do not hesitate to ask any questions related to this exercise in the "Exercises Q and A" *Microsoft Teams Ubiquitous Computing Spring 2021* channel. Make sure to not give away parts of the solution when asking a question. If you have questions that entail partial solutions or details about your chosen approach, please contact us directly via e-mail. Send your e-mail to paul.streli@inf.ethz.ch or manuel.meier@inf.ethz.ch and start the subject with *[Ubicomp21 Exercise]*.

6 Changelog

Here we will keep a changelog of this document. Corrections may be released over the course of this exercise. Please check the *Microsoft Teams* channel to stay up-to-date.

References

[1] Mojtaba Khomami Abadi, Ramanathan Subramanian, Seyed Mostafa Kia, Paolo Avesani, Ioannis Patras, and Nicu Sebe. Decaf: Meg-based multimodal database for decoding affective physiological responses. *IEEE Transactions on Affective Computing*, 6(3):209–222, 2015.