Individual Assignment for 02445

June 13, 2025

Deadline to submit: 26.06.2024; 23:59

TASK 1 (55% of the assignment score): AI Models reporting in Research Papers

The objective of this assignment is to critically analyze and compare AI research papers, focusing on the accuracy and appropriateness of the evaluation methodologies employed to communicate the efficacy of the AI systems.

Research Papers Selection

Two folders are shared with you under

```
assignment/task_1: papers_negativeExample; papers_positiveExample
```

in

```
assignment.zip
```

Each folder comprises of 5 papers, with negative and positive aspects of reporting AI models, respectively. Please choose (any) 2 papers: one from each (positive & negative) folder and describe:

1. why paper X is an example of appropriate evaluation of their AI models output.

2. how paper Y has inadequately present their AI output, either due to methodological flaws, misinterpretation of results, or inappropriate evaluation metrics.

Deliverable: Maximum 1.5 page summarization of:

1. State which is your paper X and Y, and what the papers evaluate.

2. Why the choice of evaluation is appropriate/inadequate.

3. What should have been considered in paper Y (or) how can the evaluation be made adequate?.

Note:

- The paper: *Model-cards for model reporting* can provide mode context to the importance of this task as you continue to work with AI systems.

TASK 2 (45% of the assignment score): Device a suitable cross-validation scheme

The problem comprises of predicting affect, specifically **Frustration**, from heart rate signals. All the materials are shared under:

```
assignment/task_2
```

in

`assignment.zip`

You will be using a part of the EmoPairCompete dataset [1] which is described in:

`data_description.pdf`

The data can be found in the file:

`HR_data.csv`

The variables are described below. Note that the dataset you are working with in this project is a subset of EmoPairCompete; the paper is shared as:

`emopaircompete_physiological.pdf`

Variables

Variables	Description
HR_Mean	Mean of the time-domain heart-rate signal for the individual
HR_Median	Median of the time-domain heart-rate signal for the individual
HR_std	Standard deviation of the time-domain heart-rate signal for the individual
HR_Min	Minimum of the time-domain heart-rate signal for the individual
HR_Max	Maximum of the time-domain heart-rate signal for the individual
HR_AUC	Area under the curve of the HR
Round	Puzzling round (1-4)
Phase	Phase of data collection in a round (1-3)
Individual	Index of the individual
Puzzler	Was the individual puzzling or instructing
Frustrated	Level of frustration experienced rated by the individual
Cohort	Different population groups

Task 2 deliverable: maximum 2.5 pages(including figures, tables, etc)

Use the template share for the project report. The steps are:

1. Implement two (recommended) different machine learning (ML/AI) models to predict/classify the different levels of frustration based on the HR features. In the report describe the following and the reasons behind the them:

   - Which AI models (can be ML models, eg: RF, decision-trees, logistic regressions, ANN, etc). **No extra points for the most sophisticated model.**
   - Classification or regression output.
   - Number of input features ($>= 1$)

2. **FOCUS: The data comprises of repeated measures (round, phases) from the same individuals and different cohorts. Apply a suitable hold-out or cross-validation technique and reason your choice.** Define your goal: eg: you developing a generalizable/personalizable model.

3. Compare the performances of your models using suitable statistical tests (if applicable). Some clues to answer this question:

   - Model robustness to changes in data
   - Model generalization
   - Consistency

Notes:

- The point of Task 2 is NOT to build fancy AI systems, or get 99% accuracy. The point is: Given the data, and your goal, what hold-out/CV scheme do you use.

References

[1] Sneha Das, Nicklas Leander Lund, Carlos Ramos González, and Line H Clemmensen. Emopair-compete - physiological signals dataset for emotion and frustration assessment under team and competitive behaviors. In *ICLR 2024 Workshop on Learning from Time Series For Health*, 2024.