Project: Emotional States

Report Author: Joaquin Menendez

Team members

• Sicong Zhao

• Mikella Green

**What did I accomplish this week?**

I developed a pipeline that trains a model using a Grouping Shuffle CV, stores the regression metrics and the accuracy metrics and return a data frame with all this metric for the different N folds. The pipeline binaries the regression output for every fold and calculate the scores).

I checked if we get good results transforming our regression predictions into binary labels for both valence and median valence. The metrics are not much better than chance for the Valence metrics, but we had good metrics for the median valence metric.

**How many times did you get stuck this week in some significant way?**

3. We discover problems with the data, and we do not have access to the complete brain data. Also, the Fitbit employee whom I communicate didn’t answer my message yet.

**Was there anything that you set out to accomplish this week but did not?**

I wanted to train some models using the brain data, but we only had access to the data from one experiment. We asked Mikella and she is trying to figure it out what happened

**If your project is moving slower than expected**

We need to wait for Mikella for this data. Nevertheless, we could still build a model without brain data nor survey data (only Fitbit data) to have it as a baseline.

**• How does this slow down affect your overall schedule?**

We could still advance, the person that is most affected is Mikella.

**• Are you still on track to finish your project?**

I consider that we are.

**What do you plan to accomplish in the following week?**

I am going to take a break while I wait for Mikellas’s data and, because I am traveling for my spring break. Nevertheless, I will try to run a baseline model using SVM, Bayes Ridge and RF using only FITBIT data. Also, I would like to make some modification to the CV pipeline in order to train a MERF model. (The MERF model data structure than the rest of the models in Sklearn)

**What if any, help or resources do you need from your client, project managers, or MIDS?**

Yes, I still want to meet with someone that works with heart rate and Fitbit to discuss feature engineering.

Hour card

02/02/2020 Monday

13:30 – 15 Worked developing a pipeline to train a model using a Grouping Shuffle CV, Store the regression metrics and the accuracy metrics. The pipeline binaries the regression output for every fold and calculate the scores).

15 - 16 Meeting Echelen.

19 - 24 We realized that we had some data incongruities in our dataset. Mostly given by the fact that the researchers used different scales to measure the same things. Worked solving these issues.

02/03/2020 Tuesday

14 - 15 Meeting with Greg and Mikella - We discussed some doubts about data incongruity. We collected all the information needed to solve this problem.

22-23 Finished cleaning data

02/04/2020 Wednesday

10-13 Re-run all the models with the corrected data. The results did not changed too much. The best model still being random forest.

15-16 Meeting with the groups: Conservation International and Durham County/CJRC

TOTAL 13.5 hours