Thesis Report 3 : 02 March - 09 March

Goals

- Finish extracting LMA features from all MoCap data (on the remote server) ✅
- Write META files to count which and how many emotions/lma frames we have ✅
- Start creating and training models to go form LMA features to Emotion ✅
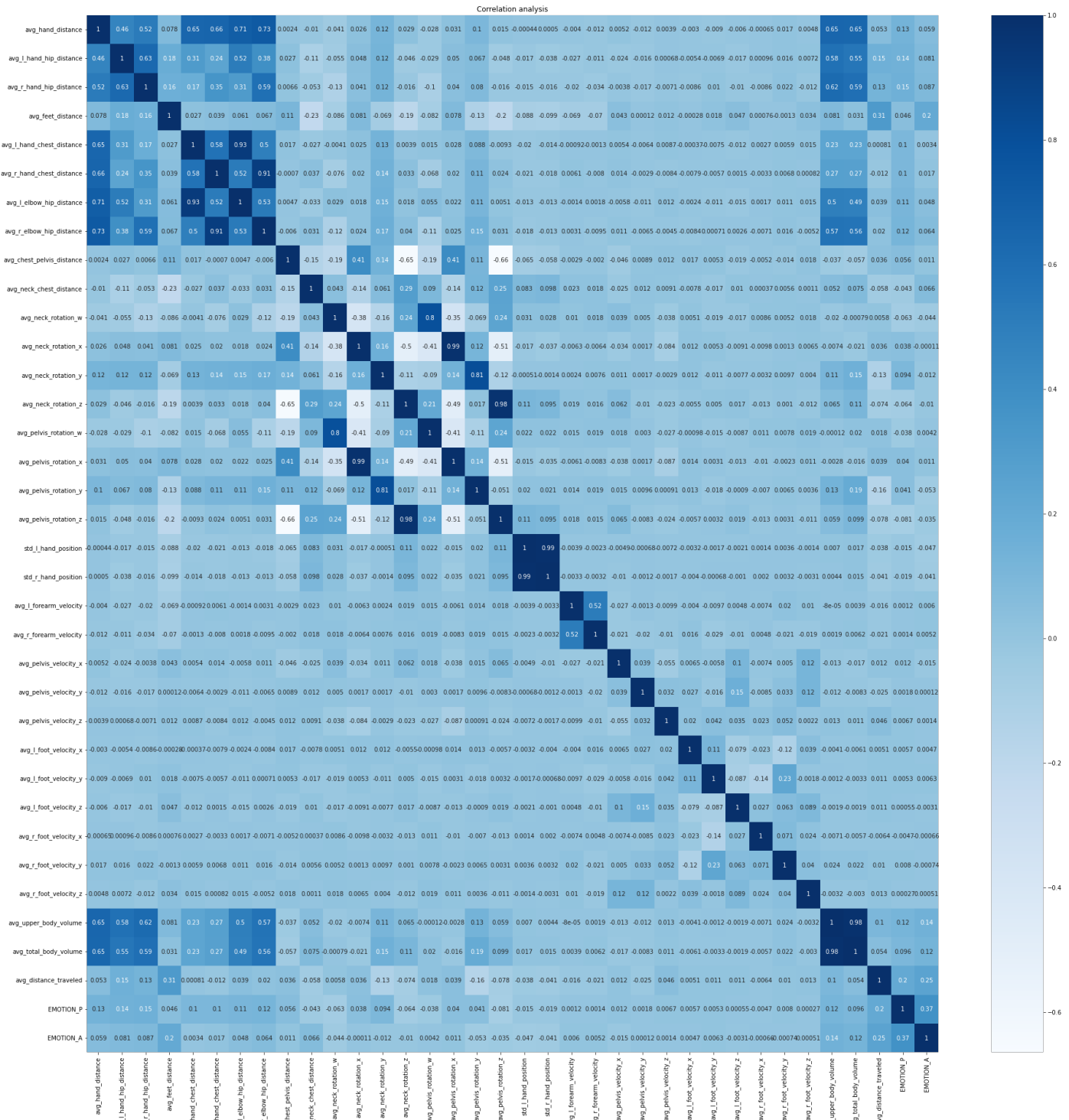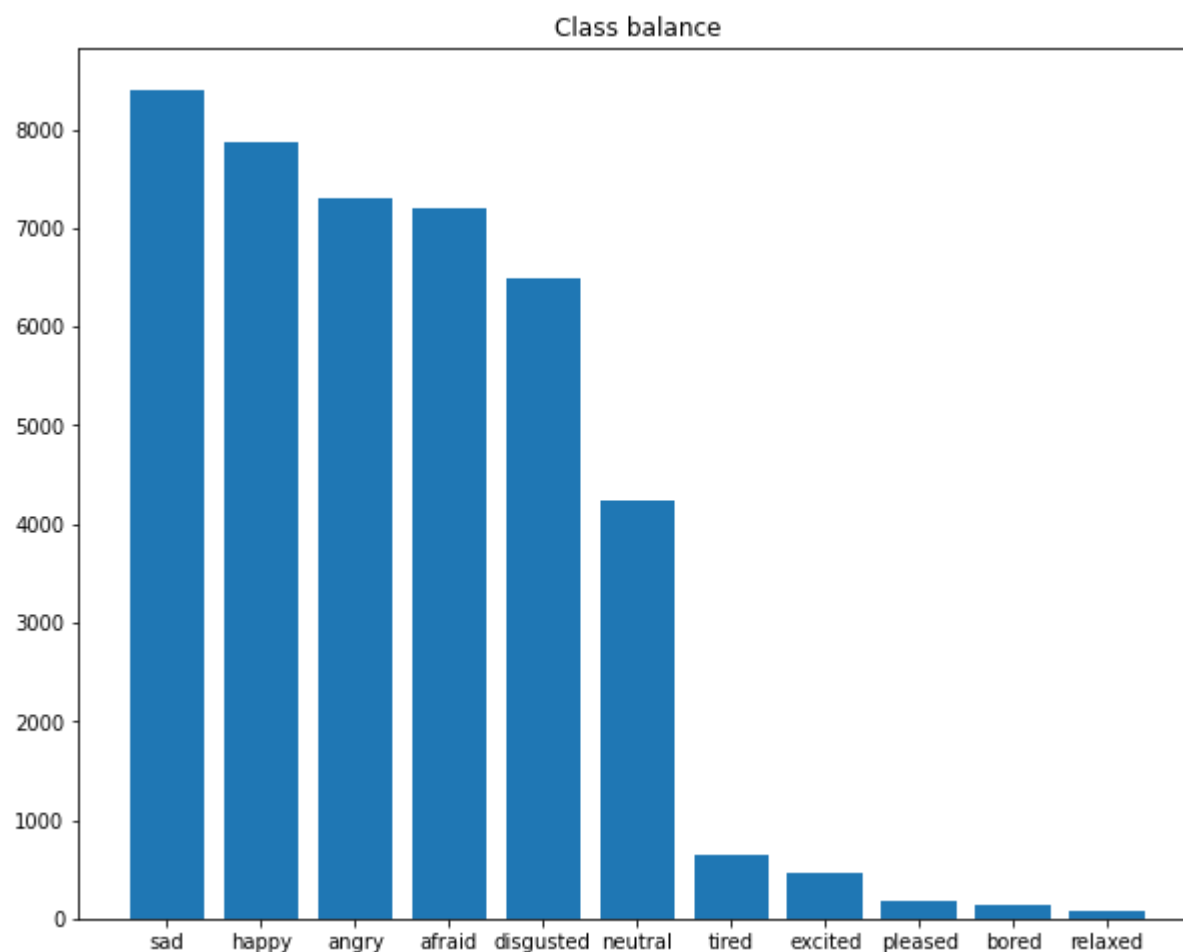
Last Week Leftovers:

None

Done

- Finished extracting LMA features from all Deepmimic converted data

- Created META files with all converted file names and emotions

  - In total we have: 43039 samples (60hz)
  - **11 Emotions** - Angry, Happy, Sad, Disgusted, Afraid, Neutral, Pleased, Bored, Tired, Relaxed, Excited

- Created Datasets by writing all LMA feature samples into CSV files

- Converted the Emotion Labels into PA coordinates (Note: may need to revise these values as they were created by "olhometro")

  - "angry": (-.5, .6)
  - "happy": (.6, .55)
  - "sad": (-.7, -.3)
  - "disgusted": (-.4, .25)
  - "afraid": (-.3, .7)
  - "neutral": (0, 0) (note: may be worth removing neutral as it may be messing up the regressions)
  - "pleased": (.65, .2)
  - "bored": (-.65, -.8)
  - "tired": (.1, -.85)
  - "relaxed": (.5, -.6)
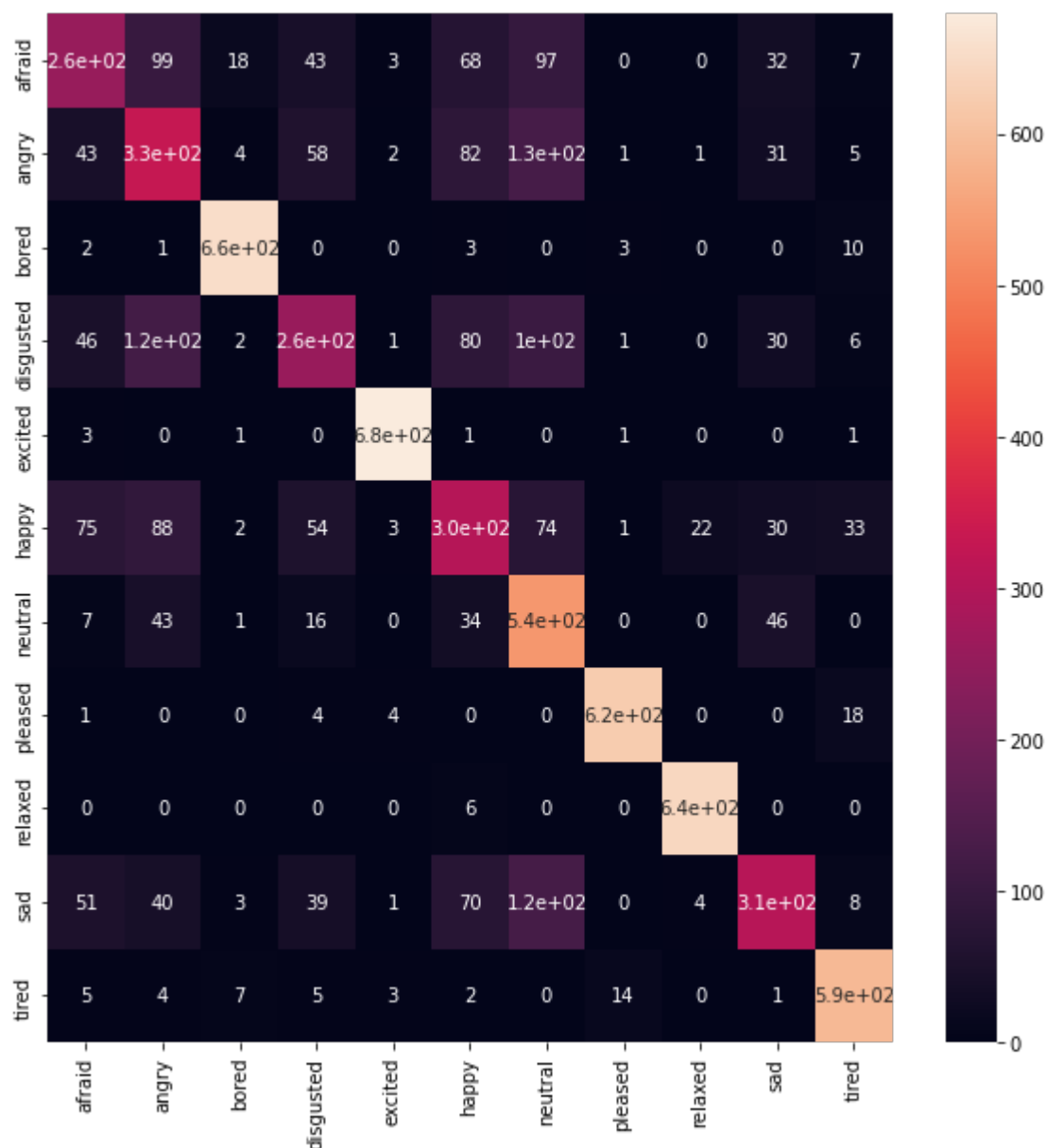
○ "excited": (.5, .8)

- **Performed Data Analysis**:

    ○ Reported on the types of variables, granularity and correlation between features

    ○ This was all done on a Jupyter Notebook

- **Performed Data Preparation**:
    ○ Created datasets with the data normalized, balanced (using both SMOTE and Undersampling) and with feature selection (by removing correlated variables)
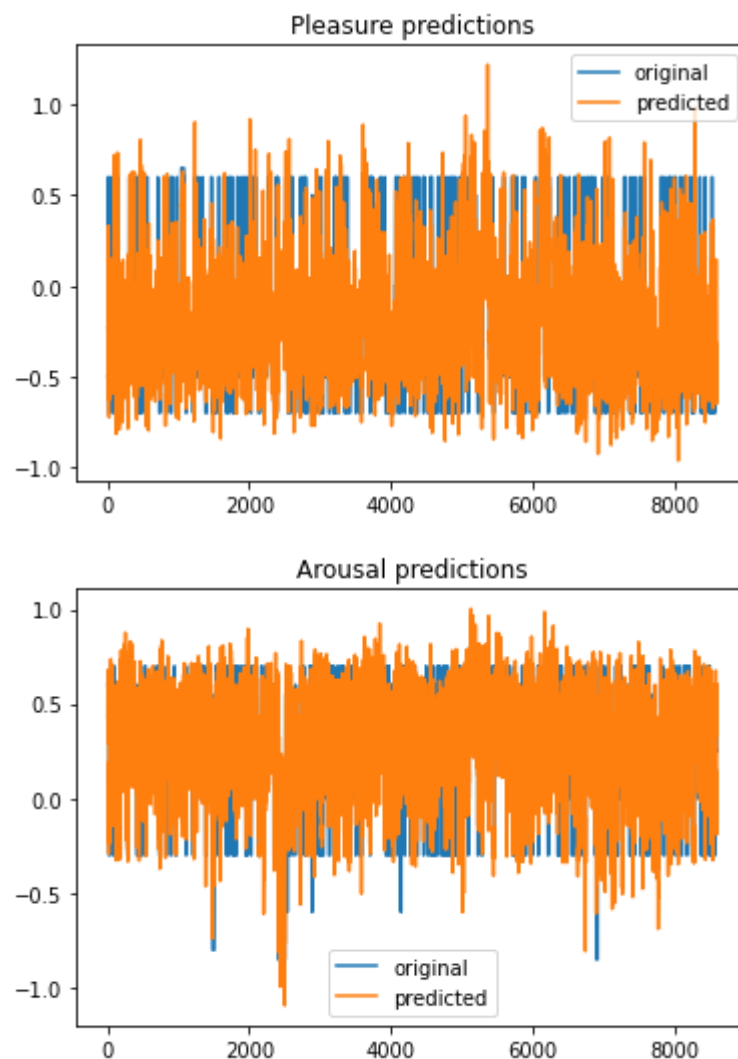
- **Created a simple MLP classifier to identify LMA features by label**:
  - Overall, despite the model being super simple and not having done any hyper parameter tweaking, we still obtained +- good results using our prepared data (around 77% accuracy)

- **Created a simple MLP network to perform a regression, mapping LMA features to PA coordinates**:

  - Overall, despite the model being super simple and not having done any hyper parameter tweaking, we still obtained +- good results, having a total 0.11 Mean Squared Error on our test set

- **Used Gradient Boosting to perform the regression from LMA features to PA coordinates**:

  - Tried utilizing SciKit learn's GradientBoostRegression for our problem. Results weren't too great, but I didn't spend too much time tweaking the hyper parameters

  - Then, I swapped over to XGB's **XGBRegression**. Results were pretty good and this may not be too bad of an avenue to pursue, should Neural Networks not have as good of outcomes as we desire

Pleasure predictions

Arousal predictions

Left Undone

Nothing

Problems

Notes

Thoughts

This month I really wanna be done with both the LMA-PAD and PAD-LMA classifiers so that I can dedicate April to motion synthesis and GUI and May to thesis writing and User testing. As such, I must admit I was a bit anxious to get into the machine learning aspect of this project. Nevertheless, and despite not having been available to work during the weekend, we still managed to reach some meaningful progress. If nothing else, we can at least be sure we can fallback and rely on simple Label classification rather than using regression.

Also, it should be noted more data preparation can be done in order to maybe improve our results.

The "Neutral" emotion may be best left unused and although we used data balancing for doing Label classification, we didn't balance our dataset for our regression, so we ended up with too many of certain emotions, with others underrepresented.

Work Hours

- Worked Friday from 12pm to 3pm

- Worked Monday from 12pm to 6pm

- Worked Tuesday from 12pm to 8pm

- Did not work during the weekend, but left remote server extracting LMA features