

1. Project Statement

The project was based around data collected during a 2012 study conducted at the University of Oslo and resulted in a paper, Jensenius et al., "The Musical Influence on People's Micromotion when Standing Still in Groups", Proceedings of the 14th Sound and Music Computing Conference (2017). The study is described as follows:

"It is commonly assumed that listening to musical sound, and particularly dance music with a clear pulse, makes us move. However, most empirical studies of music-induced motion have mainly focused on voluntary and fairly large-scale movement. This dataset was collected as part of a study which aimed to investigate the effects of music stimuli on movement when participants try to remain at rest. We collected data through optical motion capture from groups of people instructed to stand as still as possible with and without music stimuli. We then looked at the differences in movement between conditions." ¹

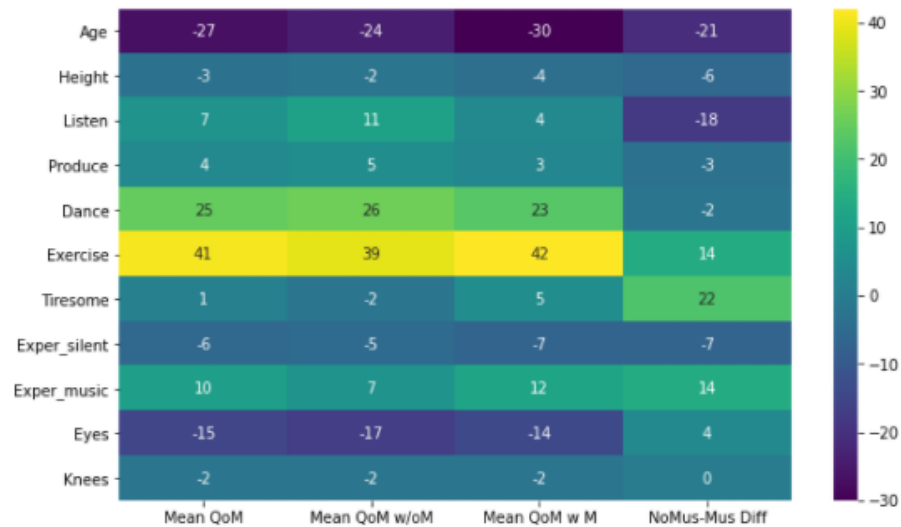
Using the data collected during the study we are looking to predict future head movement based upon silence/music at the preceding time as well as movement up until that point. If we find that there is predictive power in this previous movement then it could point to a stronger likelihood of usefulness for the field of musicology. By determining what factors appear to most be affecting unconscious movement, we could further unlock how they could affect unconscious thought. We will specifically look to predict the sum of movement over the next 5 seconds, 10 seconds and 20 seconds into the future.

2. Datasets

The collected data is represented in three categories: Motion, Music, Demographics.\

- Demographics were collected via a survey given after the experiment. For the modeling portion we removed the columns that included ex-post data to prevent the model from receiving "future" information.
- The demographics data also included quantity of motion (QoM) data for each participant for portions of the experiment with and without musical stimuli. We ran Pearson correlation that was plotted to a Heatmap to gather a general idea of what features most correlated with motion. No feature showed a correlation over 0.50 , most were much less. Exercise was the highest positive correlation at approx. 0.40 for all types of motion, while Age had the highest negative correlation with approx.. -0.28. The correlation for each relation ship is shown in figure 1 below.

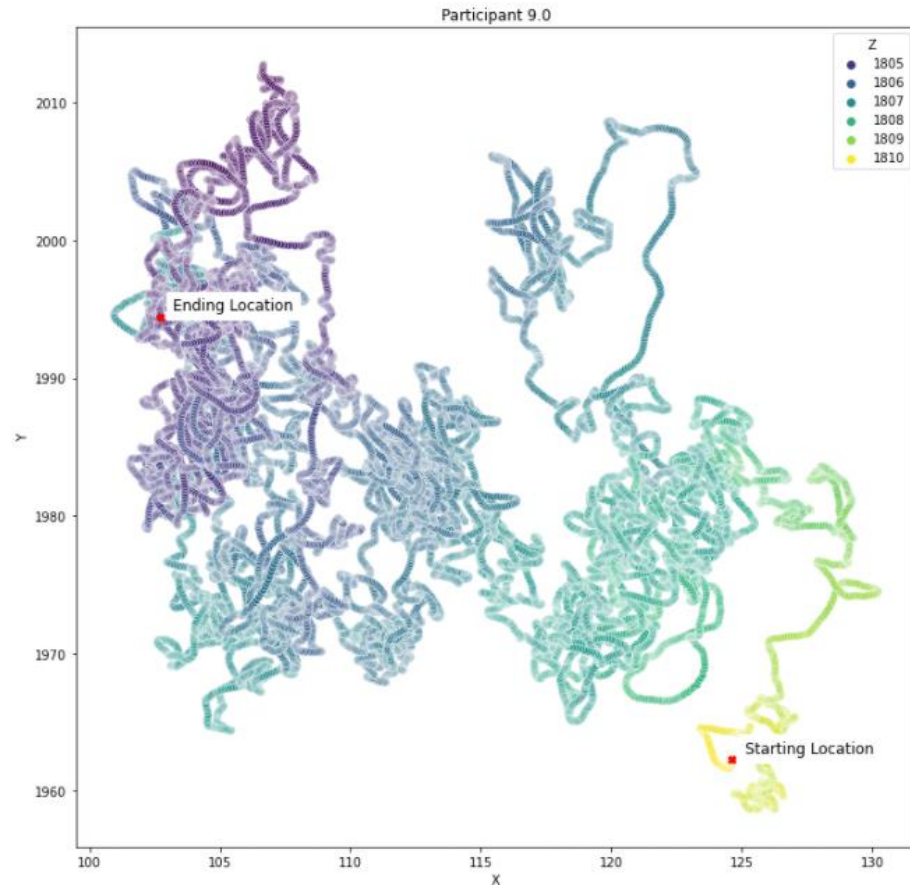
Figure 1



- Motion was collected as follows:

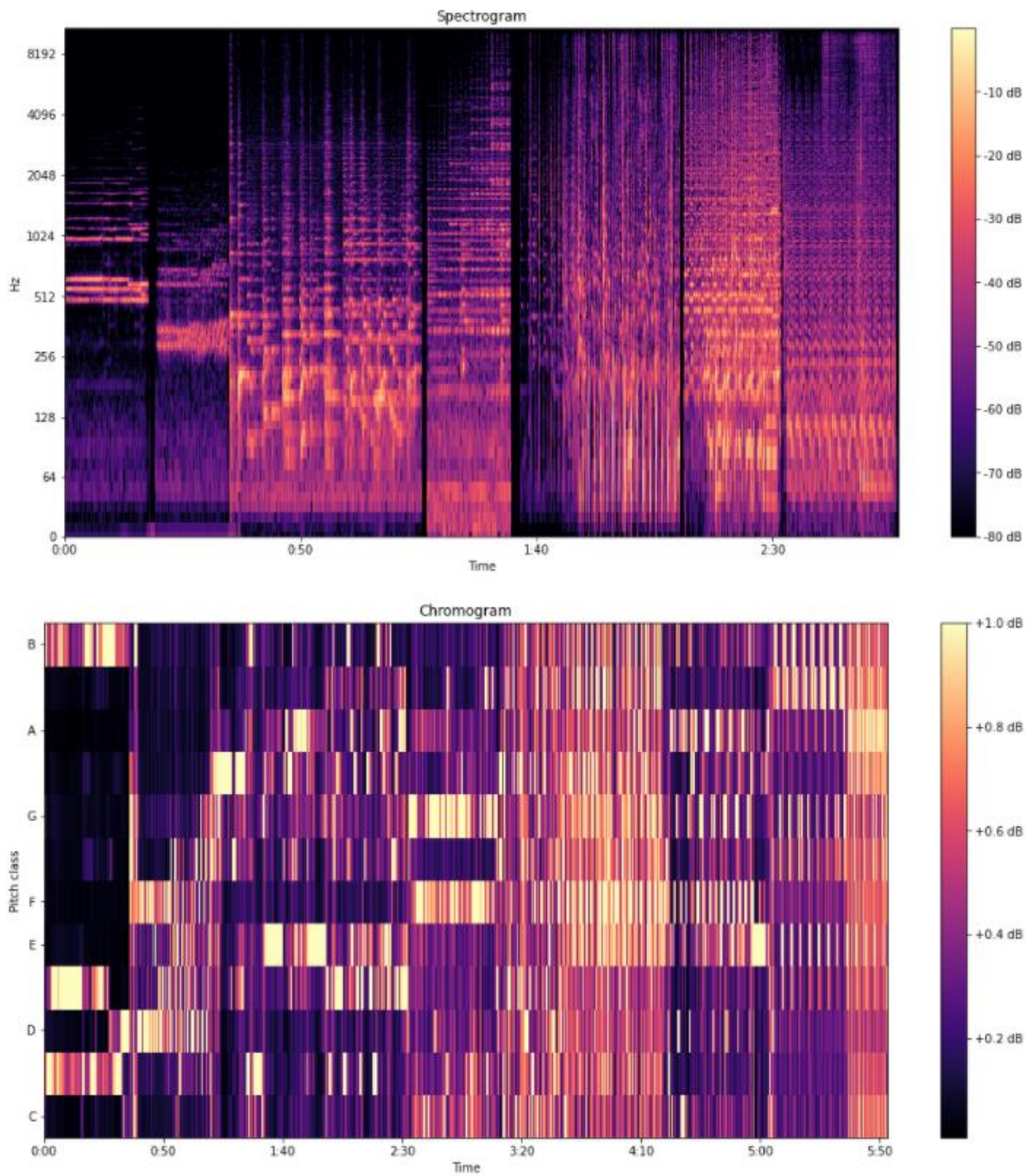
1. "The instantaneous position of a reflective marker placed on the head of each participant was recorded using a Qualisys infrared motion capture system (13 Oqus 300/500 cameras) running at 100 Hz. The data were recorded in 8 groups of 12-17 participants at a time. Participants were asked to stand as still as possible for 6 minutes, starting with 3 minutes in silence and followed by 3 minutes with music. Participants were aware that music would start after 3 minutes, and were free to choose their standing posture. The distribution of participants in the recording space was standardized across trials with marks on the floor indicating the approximate feet position. The motion capture system was triggered and stopped automatically with the stimuli playback system, thus all recordings are exactly 6 minutes long."¹
2. An example of the collected data is shown in Figure 2 below. The visual represents a view looking down onto the top of a participants head with the following corresponding planes
 - X – left to right, or side to side
 - Y – front to back
 - Z – height represented by color with lighter being higher and darker being lower

Figure 2



- Music used for stimuli during the later 3 minutes was the following¹:
 1. Lento (#3) from György Ligeti Ten Pieces for Wind Quintet (20s)
 2. Allegro con delicatezza (#8) from György Ligeti Ten Pieces for Wind Quintet (15s)
 3. Adagio from Joaquin Rodrigo's Concierto de Aranjuez (40s)
 4. Winter movement from Vivaldi's The Four Seasons (20s)
 5. Left & Right by D'Angelo, featuring Method Man & Redman (35s)
 6. Marcando la distancia by Manolito y su trabuco(20s)
 7. Cubic by 808 State (30s)
- For purposes of model building we converted the model into a chromogram using the python Librosa package. A chromogram is a representation of the strength of musical notes (12) over time periods where a spectrogram represents frequency in hertz. The following spectrogram and chromogram represents the 3 minute portion of the experiment where the music stimuli was present.

Figure 3/4



3. Cleaning and Processing

a. Demographics

- i. The data was relatively clean with most columns being numeric type except for Group and Sex. The group column was used to create a subject ID column ('SID') column that retained the individual group divisions while a Participant ID column ('PID') was created to give a global designation. There were 17 participants who have no value for 'Eyes open?' and 'Locked knees'. This happened for all participants in the P group. We assigned 0.5, which represents both, for both of these columns. There was also one participant who didn't give an experience of motion answer, so we used a mean of the column to fill.

b. Motion

1. We added features that we can use for our model. We are going to predict the absolute value sum of euclidean motion (combination of all 3 directions) 1 second into the future. We will be saying the participant will move X mm in the next 1 second based on how they have moved in the previous time periods. The features we are creating will be made up of these past time periods and the features will be information about the time periods (min motion, max motion, average motion, etc.) that we can use to describe what happened during that time period. These features use data from the individual participants so we created them prior to joining all of the data together.

- `_disp` - absolute value of displacement per step in noted direction
- `_disp_total` - sum of disp from start in noted direction
- `step_eucl` - Euclidean distance moved in time step (combination of motion in all directions)
- `total_eucl` - Total Euclidean distance move since start (combination of motion in all directions)
- `target_1_sec` - Target for prediction, sum of motion for 1 second into the future using euclidean distance
- `*sec_hist*_` - description of 1 second starting at time noted using function noted
 - Ex. `_4sec_hist_mean_` - The mean motion in a tenth of a second during the time period from 4 seconds previous to 5 seconds previous
- Aggregating functions currently used are mean, min, max and standard deviation

c. Music

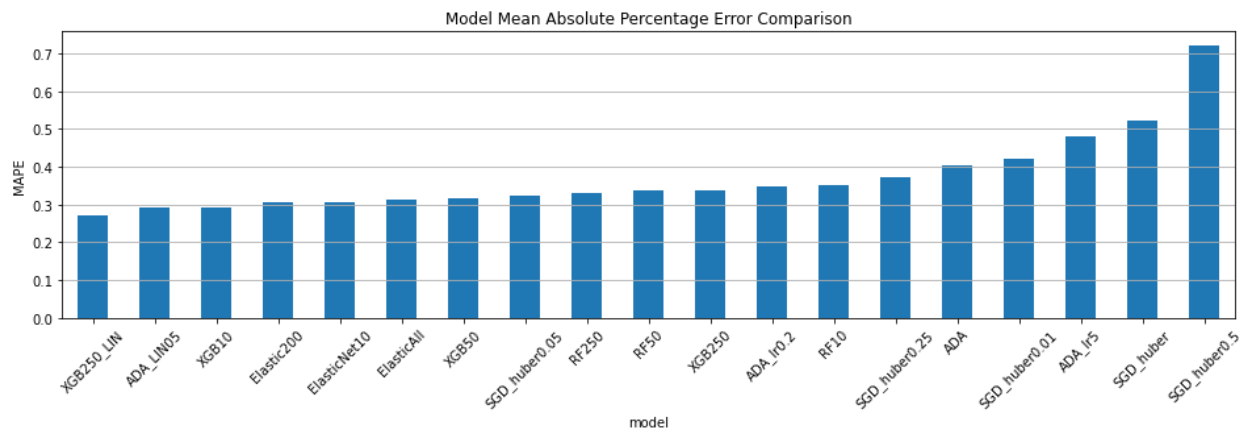
- i. Similar to how we added new features for motion, we added features for the past amplitude of music notes. The features we created were made up of these past time periods and the features were information about the time periods (min note amplitude, max note amplitude, average note amplitude, etc.) that we can use to describe what happened during that time period. As previously mentioned the music was pulled in based on its note values for the 12 common notes C, C#, D, D#, E, F, F#, G, G#, A, A#, B. Features were created for each note.

d. Resampling

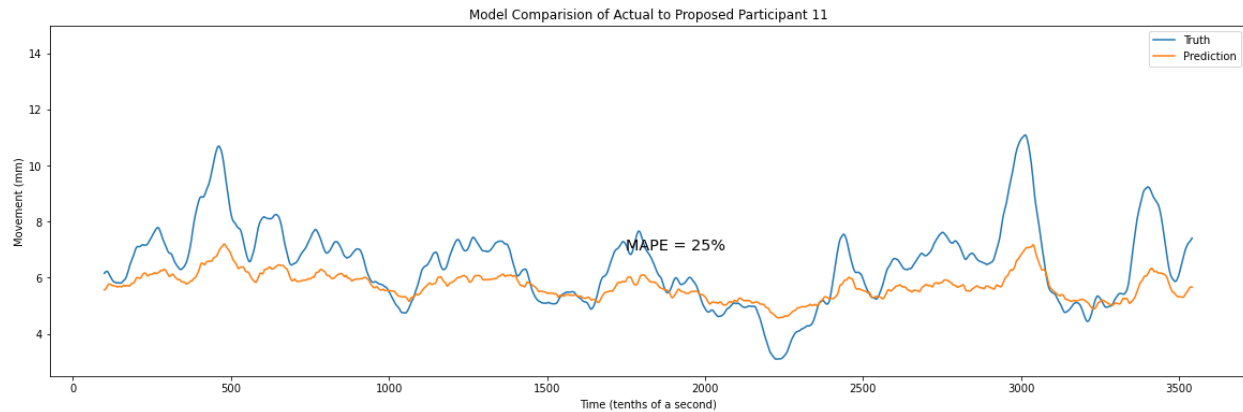
- i. The music data was brought in at 22 Hz or 22000 cycle per second. The motion data was captured at 100 readings per second. In order bring them into a single dataframe as well as with thought to memory/computation restraints, the music and motion data were both sampled to one reading every 1 tenth of a second. The mean method was used during resampling. This resulted in just under 3600 rows of data for each participant and over 250,000 rows of data during musical stimuli in total.

4. Modeling

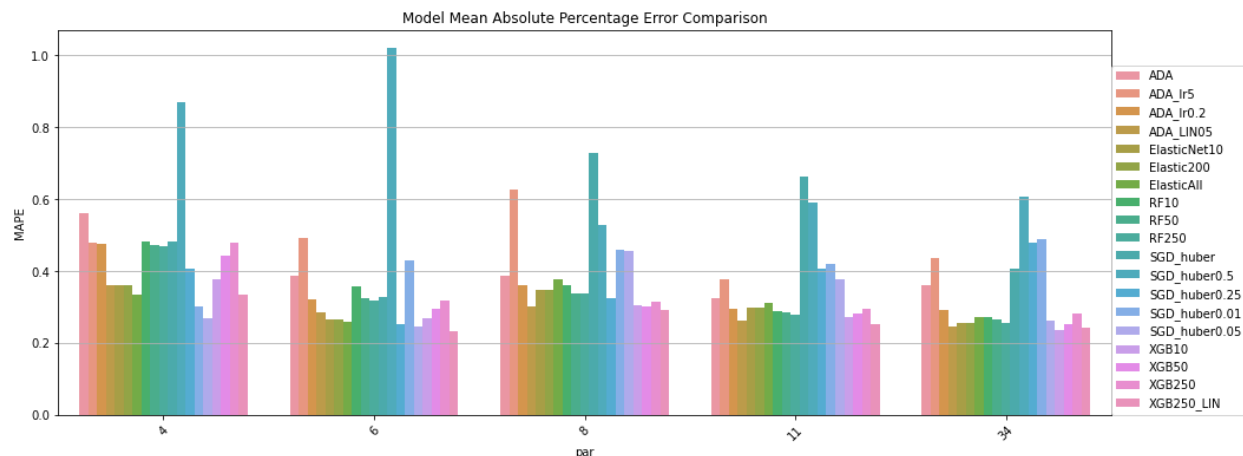
We fit a number of different models to the data both tree based and linear. We found that the linear models worked better at achieving a low error rate for our metric mean absolute percentage error. Tree based methods did do a better job of being sensitive to the amplitude of the motion, but not in a consistent enough manner to achieve better metrics. Below you can see the final scores for all of the models. Each model was trained on 10 random participants and then tested on 5 different random participants. The results in the visual below are averages of the 5 participants.



XGBoost Regressor has the best overall performance and also was one of the more consistent performers. Below is a visual of the predicted motion (orange) vs. actual motion (blue). While the model prediction hugs the mean much more closely than the true values, it does consistently trend in the correct direction. This means it would be useful in predicting someone reaction to musical stimuli as far as increasing or decreasing, but not the actual amount. It is a good first step.



The below visual shows each participant and each model that their data was tested with. The most informative item from this visual is that it shows some participants are consistently harder to predict than others. Nearly all models performed worse on participant than they did on 11 and 34. This could point to someone who isn't a good candidate for musicology or just someone that a musicological treatment has been devised for yet. Further study different participants and increasing the training set could help with different participants as could splitting different kinds of participants into different models: high motion vs. low motion, men vs. women, old vs. young, tall vs. short, quick change in motion vs. slow change in motion, etc.. There are many avenues that can still be reviewed with this dataset which could help unlock additional information about how musical stimuli affects us.



¹[MICRO Motion capture data from groups of participants standing still to auditory stimuli \(2012\) v1.0 \(physionet.org\)](https://physionet.org)