# Predicting Imagined Meters in Musical Patterns from MEG Data

Ben Limonchik, Zoe-Alanah Robert, Aashna Shroff
Computer Science, Stanford University

## Motivation

Musical data is often interpreted within a metrical framework that integrates hierarchical timing information. Listening to metronomes and imaging different meters evoke auditory responses in temporal lobe and motor-related brain areas. We aim **to use MEG data of imagined meters to classify the metrical framework (i.e. march, waltz or hemiola) that a person is imagining** while listening to a constant metronome.

## Data and Experiment

12 musicians studied. A constant metronome of 12 unaccented clicks was played with an interval of 390 ms. Participants were asked to imagine one of the following musical rhythms:

**March**: 1 2 1 2 1 2 1 2 1 2 1 2
**Waltz**: 1 2 3 1 2 3 1 2 3 1 2 3
**Hemiola**: 1 2 3 1 2 3 1 2 1 2 1 2

| | |
|---|---|
| No. of participants | 12 |
| No. of trials/ participant | 190 |
| No. of time intervals | 94 |
| No. of brain voxels | 12,180 |
| Labeled classes | 9 |

**Class Labels:**
'PD2_H32' 'D2_T2' 'D3_T3'
'D2_H32' 'U2_T2 'M3_T3'
'PU3_H32' 'U3_T3
'U3_H32'

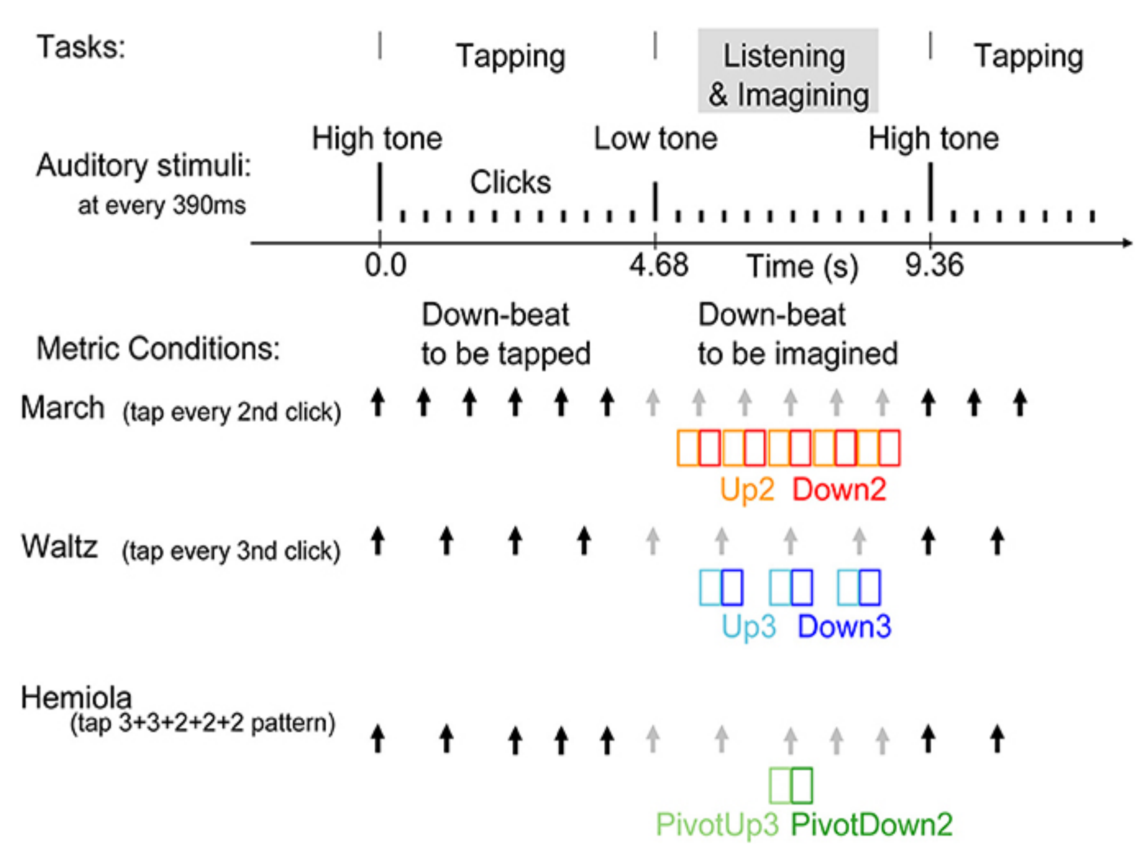
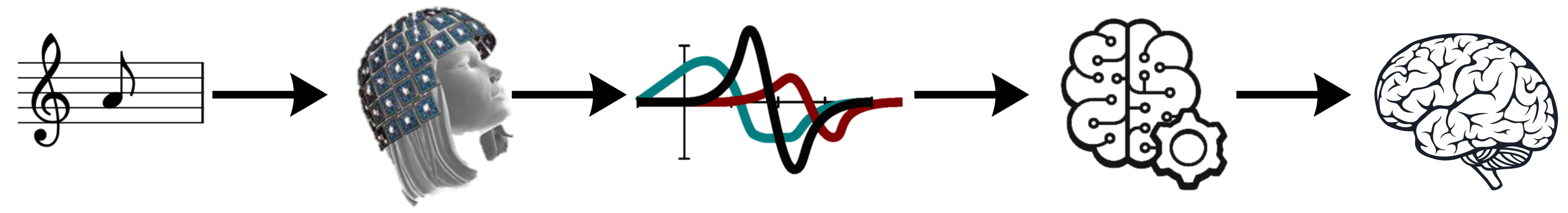


Figure 1: Stimulus sequence and tasks

Overall, the data consisted 190 trials * 12 participants * 9 classifications = **20,520 data points**, each with 12,180 voxels * 94 time-intervals = **1,169,280 features.**

## Methodology




## Feature Engineering

Given the large number of observations and features, it was necessary to perform feature reduction to save storage and computation time and increase comprehensibility.

- Any voxel outside the brain periphery with a zero-reading was removed, resulting in 4,840 remaining cubes.
- The first and last 8 MEG measurements in the time-frame were removed so only 78 time-intervals remained.
- Brain voxels were spatially averaged according to the Talairach coordinate system. Each voxel was mapped to one of 72 well-known regions of the brain.
- Features were then selected using a filter approach. Applying the ttest on each feature and comparing the *p*-values, it is seen that 75% of the 5616 (72 avg. cortextes x 78 time intervals) had strong discriminative power. However, as seen in Figure 5, over-fitting occurs when more than 50 features are used. Best 50 features were selected according to their *p*-values.

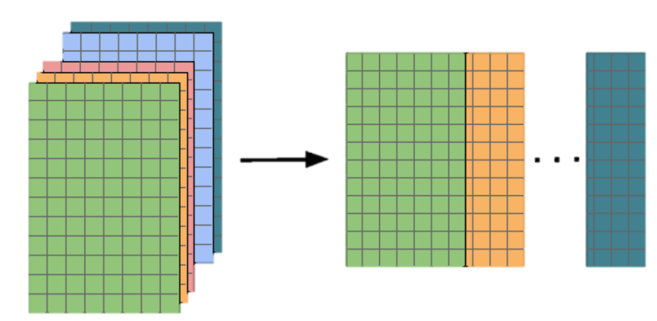


**Figure 2:** Flattening of 3D spatial-temporal tensor into 2D matrix. Each row contains brain signals across all time intervals.

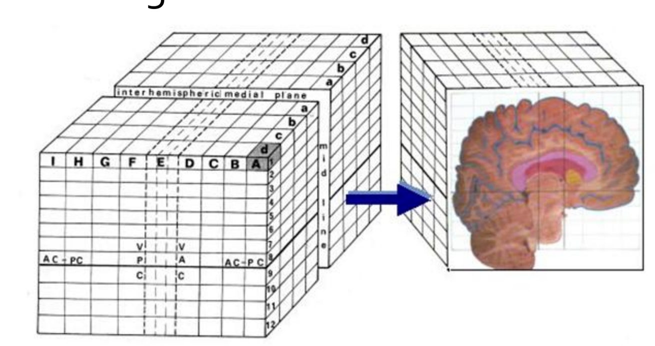


**Figure 3:** Human brain with superimposed Talairach grid

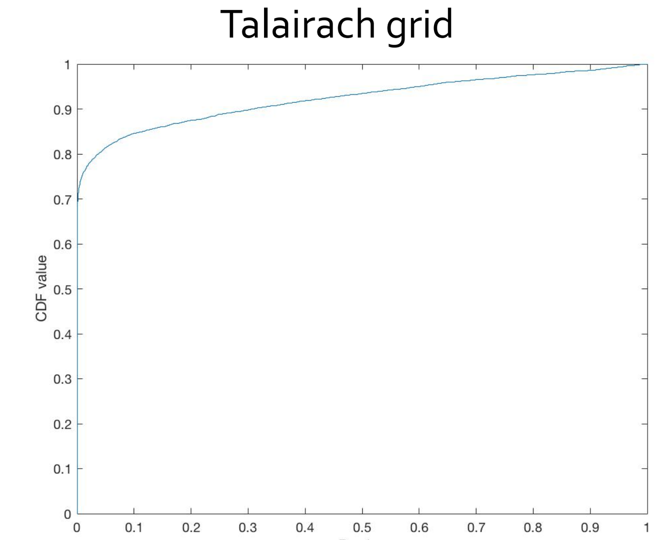


**Figure 4:** Empirical CDF of *p*-values

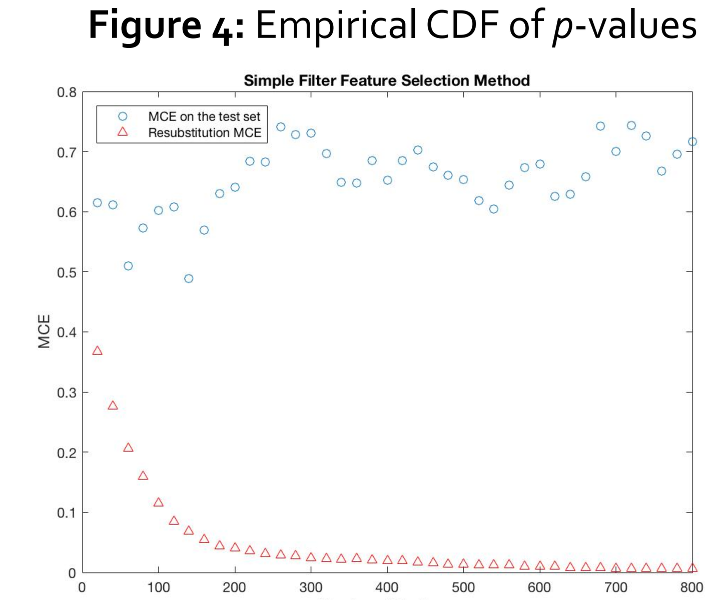


**Figure 5:** Resubstitution and Test Misclassification error for different features

## Results

Both **binary classification** between 'D2_T2' and 'PD2_H32' and **multiclass classification** between 'D2_T2', 'D3_T3' and 'PD2_H32' was performed. Model was trained using trials of 10 people, and tested in two ways:

- Using **'seen people'**, which the model was trained on
- Using **'unseen people'**, which model was *not* trained on

Different models were tried, including Logistic Regression, Linear Discriminant, Quadratic Discriminant and different SVMs. Overall, Cubic SVM performed better.

| Test Accuracy for | Seen | Unseen |
|---|---|---|
| **100 random features out of 377,520** (78 time intervals x 4848 non-zero brain voxels) | 99.5% | 33.33% |
| **5616 features** (78 time intervals x 72 averaged brain cortexes) | 98.2% | 38.50% |
| **50 best features** using filter approach | | 43.23% |

**Figure 6:** Multiclass classification using different feature subsets using Cubic SVM

## Analysis and Future Work

The 50 best features were mapped back to their original spatial and temporal locations and analyzed:

- **Spatial Brain Locations:** Temporal lobe and pre-motor cortex areas were common, which are known to be involved in **primary auditory perception.**
- **Temporal Locations:** Most indicative time intervals were about 0.05 s before and after the auditory stimulus.

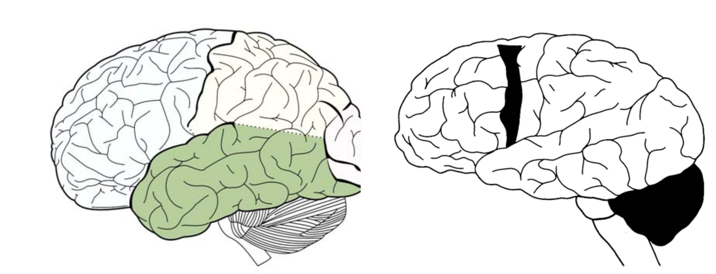


**Figure 7:** Temporal Lobe; Pre-Motor Cortex

In the future, it would be interesting to leverage feature engineering to avoid overfitting and be able to make better predictions for unseen people across more metrical classifications.

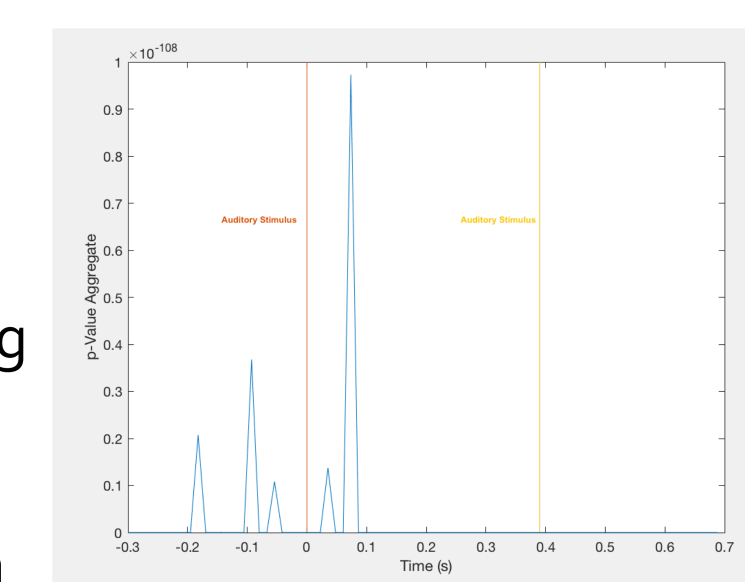


**Figure 8:** Time Interval vs. Aggregated p-values