

# Iowa Initiative for Artificial Intelligence

## Final Report

Project title:	Characterizing Naturalistic Sound Features relevant for Mental Disorders in Human Behavior and Brain Function with Computational Models		
Principal Investigator:	Dorit Kliemann, James Traer		
Prepared by (IIAI):	Yanan Liu		
Other investigators:			
Date:			
Were specific aims fulfilled:	Yes		
Readiness for extramural proposal?	Yes		
	If yes ... Planned submission date	Submitted, now working on resubmission for March 2023	
	Funding agency	NIH	
	Grant mechanism	R01	
	If no ... Why not? What went wrong?		

### **Brief summary of accomplished results:**

We have developed and built two pipelines for Aim 1 and Aim 2. For Aim 1, ridge model was used to predict brain data from the relevant sound features. For Aim 2, random forest model was used to classify Autism Spectrum Disorders (ASD) and the neurotypical control participants group.

### **Research report:**

#### **Aims (provided by PI):**

Specific Aim 1: To identify natural sound features that robustly predict functional brain responses in healthy adults.

Specific Aim 2: To discover features in natural sounds that predict ASD diagnosis from functional brain responses.

#### **Data:**

We have fMRI data of 50 ASD and matched (age, sex, IQ) 90 neurotypical (NT) participants previously collected in a large multi-site study in which participants watched 6 different short movie clips in the scanner (Table 1). The sound data are multidimensional acoustic features extracted from the movie soundtracks including: (A) a baseline set of (~10) standard audio features (incl. spectral power, the presence of speech, pitch chroma, etc.); (B) features selected to imply source proximity (e.g., reverberation (Traer & McDermott, 2016; Traer et al., 2020)) or rapidly approaching sources; (C) features selected to relate to emotional content in voices (e.g., modulation spectra (Arnal et al., 2015)). While features A

will be extracted with standard tools (librosa: <https://librosa.org/doc/main/feature.html>, pliers: <https://github.com/PsychinformaticsLab/pliers>), we have created our own tools for extracting features B and C. The combined set (A, B and C) will consist of approximately 30 features. Each feature will be extracted from the movie soundtrack once every 0.72 seconds, equivalent to the frequency at which fMRI images are acquired in our data set (repetition time, TR; 6302 images per participant). The brain data are changes in blood oxygen dependent signal (BOLD) as participants watched the movies in an 3T fMRI scanner. Thus, we have a full set of acoustic features for every BOLD data point (see Table 1). The BOLD time series data has 90k voxels parcellated into functional regions at multiple levels of granularity: 264 parcels.

<b>Movie</b>	<b>Duration (m)</b>	<b>Data points (TR = 0.72s)</b>
Partly Cloudy	5.20	444
The Office Episode 5	21	1750
The Office Episode 6	21	1750
Movie Trailers A	12.54	1070
Movie Trailers B	13.28	622
Bang! You're dead	8	666

**Table 1. Movie lengths and number of data points. Abbreviations: m, minutes; TR, MR data repetition time.**

#### **AI/ML Approach:**

In this study, two supervised machine learning algorithms were implemented using Python. For Aim 1, ridge regression was implemented to predict brain response from audio features. For Aim 2, as many extracted features may be noisy, or highly correlated with each other, Random Forest (RF) algorithm was selected to classify ASD/control group.

#### **Experimental methods, validation approach:**

Fit a ridge model with audio features.

In this pipeline, we modeled the fMRI responses with audio features extracted from the movie stimulus. The model is a regularized linear regression model. This pipeline reproduces part of the analysis described in Nishimoto et al (2011).

We first concatenated the features with multiple delays, to account for the slow hemodynamic response. A linear regression model then weights each delayed feature with a different weight, to build a predictive model of BOLD activity. Again, the linear regression is regularized to improve robustness to correlated features and to improve generalization. The optimal regularization hyperparameter was selected independently on each voxel over a grid-search with cross-validation. Finally, the model generalization performance is evaluated on a held-out test set, comparing the model predictions with the ground-truth fMRI responses.

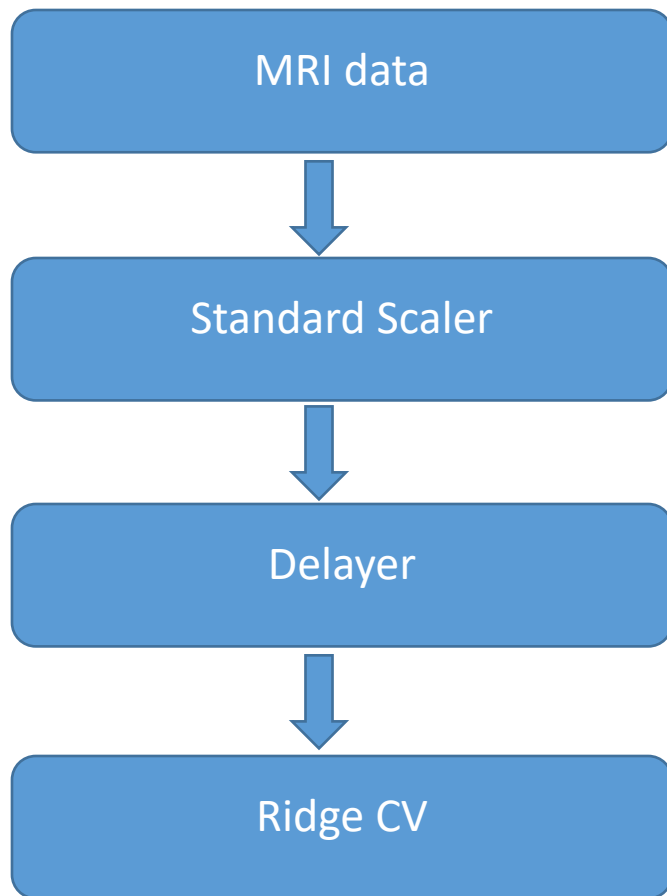


Figure 1. Pipeline for Ridge model

Fit a random forest model with BOLD activity.

In this pipeline, we classified ASD/control with BOLD features extracted from fMRI data. The model is a classification model. All missing data was replaced with -9999 and 5-fold cross validation was used for RF model due to small data set.

**Results:**

The pipelines were developed/tested and provided to the team who employed them directly to fulfil the stated aims.

**Ideas/aims for future extramural project:**

We have already submitted a funding proposal based on this initial project in November 2021 ("Characterizing Naturalistic Sound Features Predictive of Mental Disorders in Human Brain Function

with Computational Models and Multimodal Neuroimaging”) to the INI Research programs of excellence funding call. The project would extend this IIAI Pilot project by i) including additional neuroimaging data types (ECog, EEG) and ii) the application of encoding models to a different psychiatric disorder. While the initial submission did not advance, we are working towards a submission of an R01 in 2023.

**Publications resulting from project:**

Written manuscripts currently in preparation.

Invited talk to the Computational Psychiatry Workshop Ulowa April 2022.

**Reference:**

1. Traer, J., & McDermott, J. H. (2016). Statistics of natural reverberation enable perceptual separation of sound and space. *Proceedings of the National Academy of Sciences*, 113(48), E7856-E7865.
2. Traer, J., Norman-Haignere, S. V., & McDermott, J. H. (2020). Causal inference in environmental sound recognition. *BioRxiv*. (under review in *Cognition*)
3. Arnal, L.H., Flinker, A., Kleinschmidt, A., Giraud, A.L. and Poeppel, D., 2015. Human screams occupy a privileged niche in the communication soundscape. *Current Biology*, 25(15), pp.2051-2056.
4. Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology*, 21(19), 1641-1646.