Joint Feature Extraction from Functional Connectivity Graphs with Multi-task Feature Learning

Andre Altmann[1], Bernard Ng[1]
1. Functional Imaging in Neuropsychiatric Disorders (FIND) Lab, Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, USA

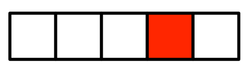
Introduction

Using sparse regularization in classifier learning is an appealing strategy to locate relevant brain regions and connections between regions within high-dimensional brain imaging data. A major drawback of sparse classifier learning is the lack of stability to data perturbations, which leads to different sets of features being selected. Here, we propose to use multi-task feature learning (MFL) to generate sparse and stable classifiers. In classification experiments on functional connectivity estimated from resting state functional magnetic resonance imaging (fMRI), we show that MFL more consistently selects the same connections across bootstrap samples and provides more interpretable models in multiclass settings than standard sparse classifiers, while achieving similar classification performance.

Methods

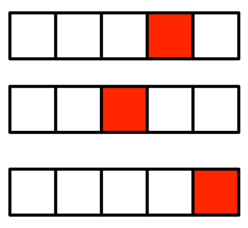
Sparse regression.

$$\min_{w} \sum_{i=1}^{N} \mathcal{L}(w \cdot x_i, y_i) + \lambda \|w\|_1$$



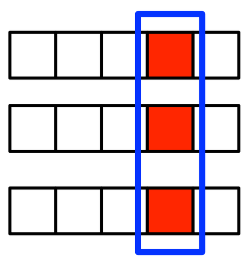
Sparse regression with *K* tasks.

$$\min_{W} \sum_{k=1}^{K} \sum_{i=1}^{N_k} \mathcal{L}(w^k \cdot x_i, y_i) + \lambda \|w^k\|_1$$



Multi-task feature learning (MFL).

$$\min_{W} \sum_{k=1}^{K} \sum_{i=1}^{N_k} \mathcal{L}(w^k \cdot x_i, y_i) + \lambda \|W\|_{1,2}$$



Data. We used rs-fMRI data from Shirer et al. (2012). Four different subject driven cognitive tasks:

- Resting state
- Episodic memory
- Music and lyrics
- Counting

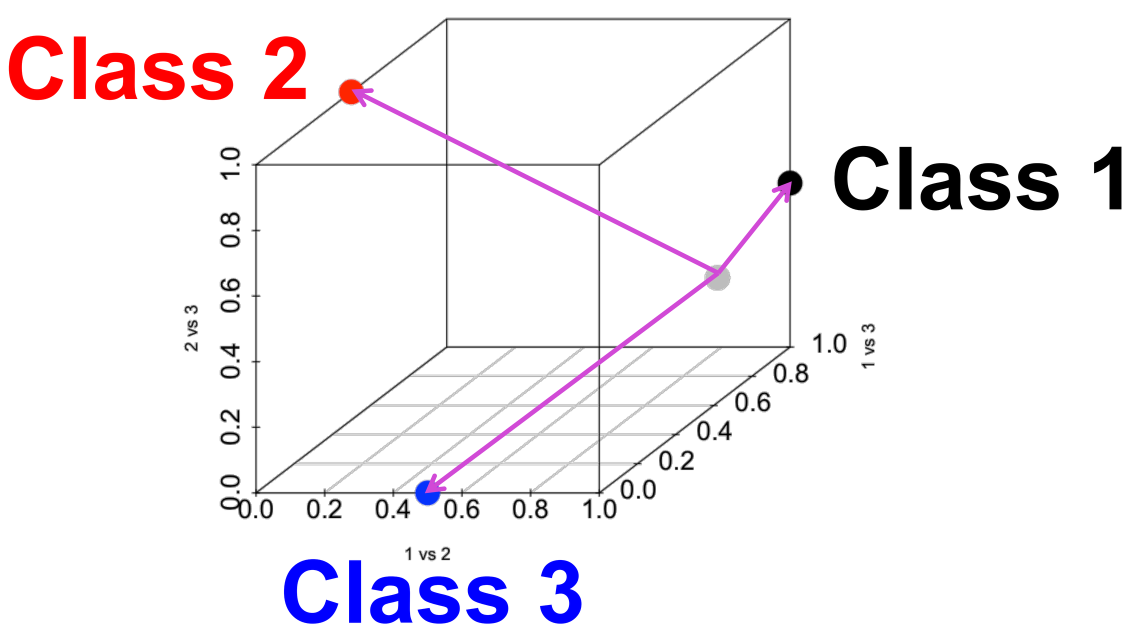
Recorded for 10 min each in all subjects (TR=2s).

24 subjects resulting in 96 data points for training.

Processing. Motion correction, normalization to MNI, spatial smoothing and HP filter at 0.008 Hz done in FSL. Confounds (WM, CSF, global signal, heart beat and respiration rate) regressed out at voxel level.

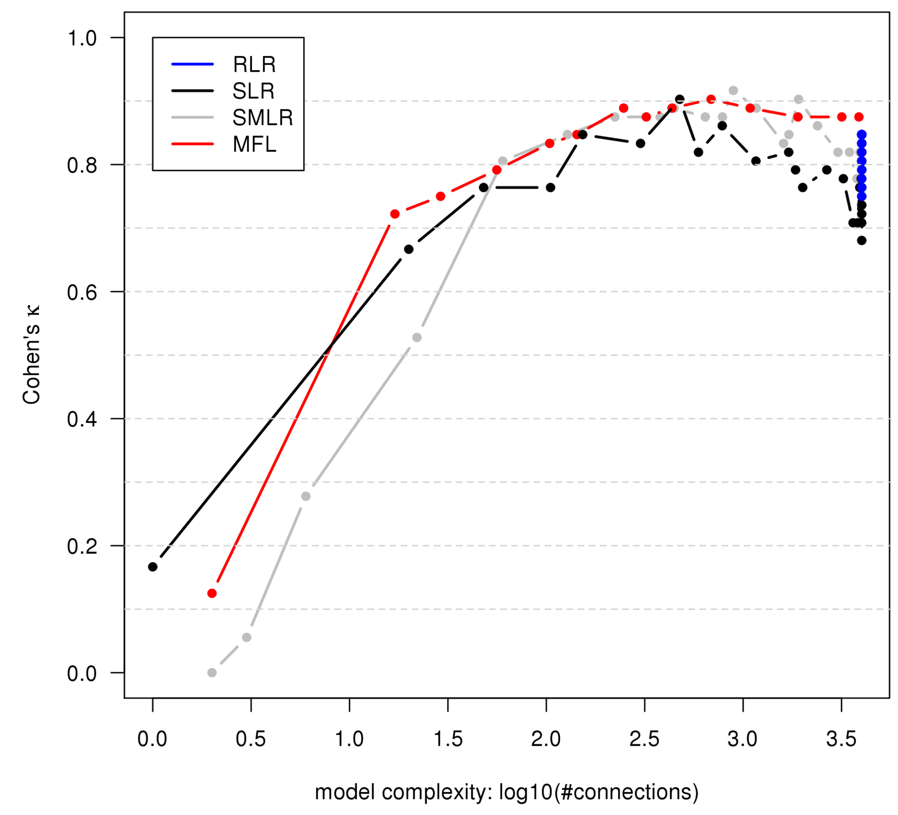
Feature extraction. Extracted time series for 90 ROIs (Shirer et al., 2012). Pairwise Pearson's corrleation converted to z-scores. Upper triangle of the resulting 90x90 matrix used as feature vector (4,005 features).

ECOC (Error Correcting Output Codes). Example with three classes.

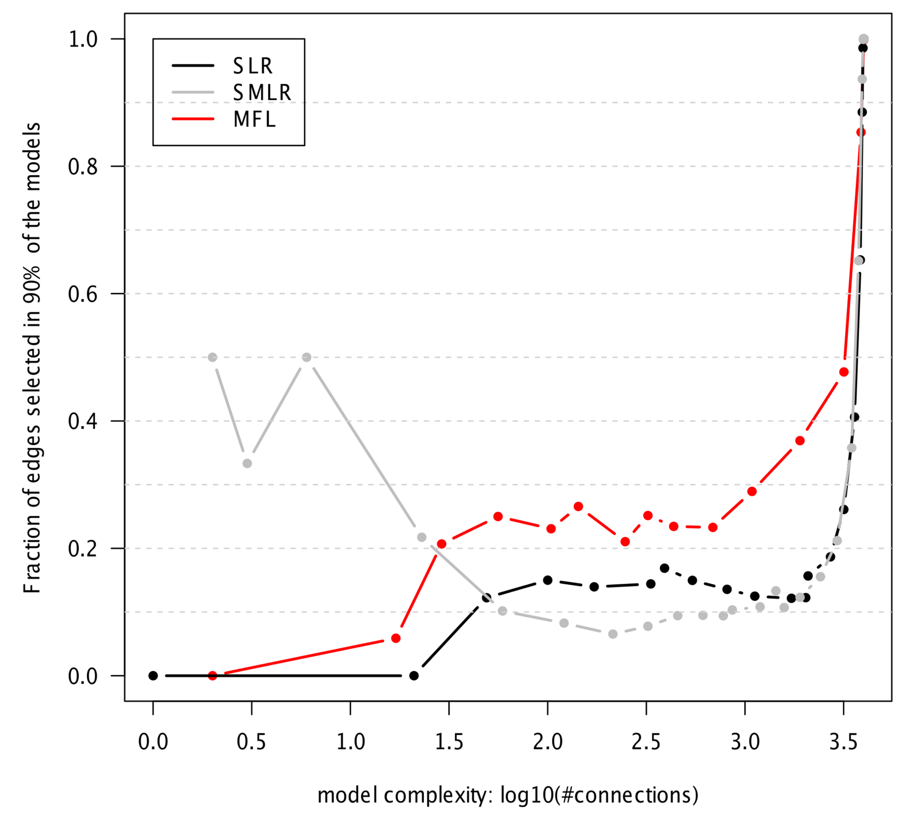


Results

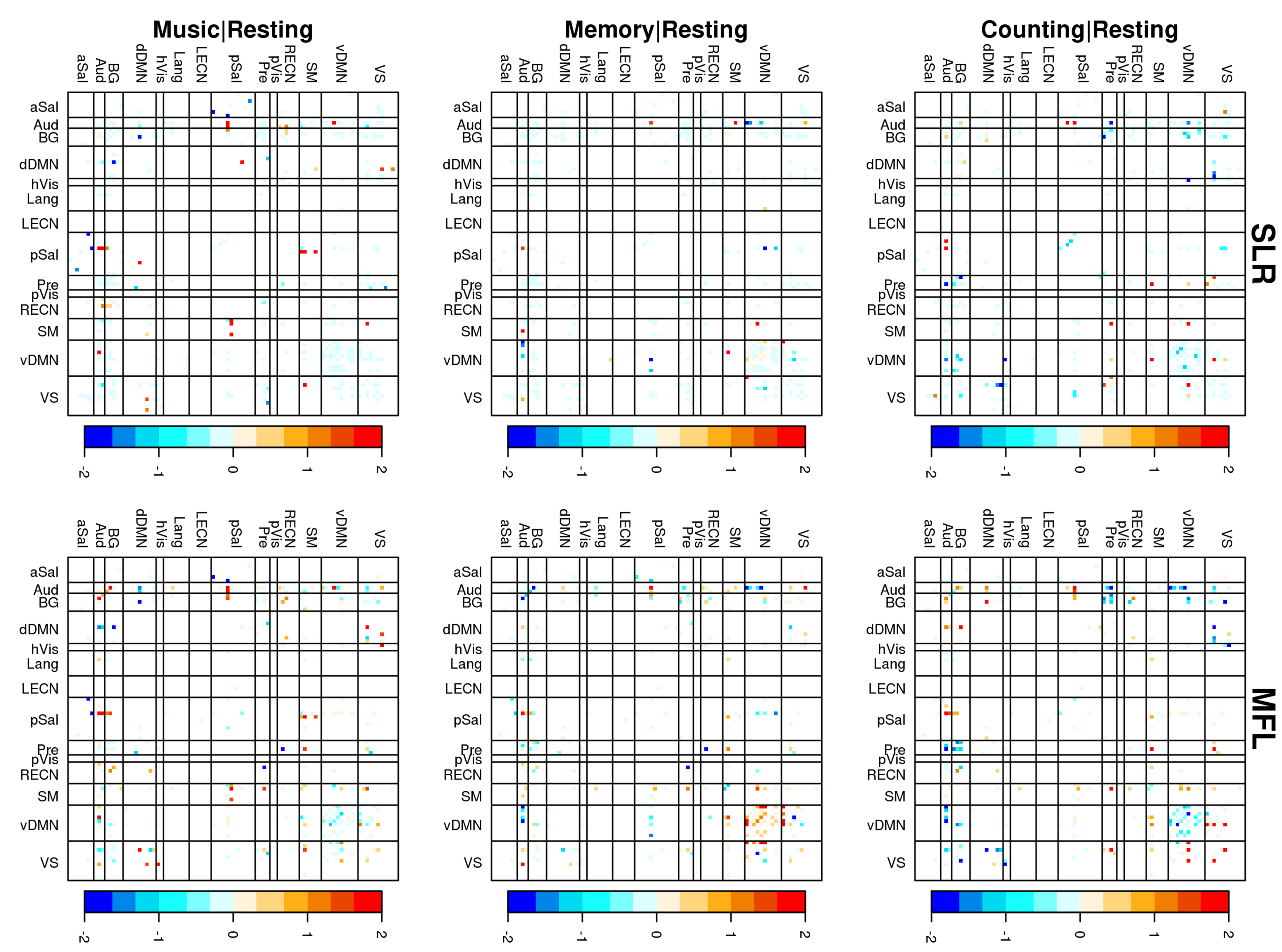
LOSO-CV performance. Comparison of MFL to reference methods: ridge regression(RLR), sparse logistic regression (SLR), and sparse multinomial logistic regression (SMLR). Methods do not differ substantially in performance wrt. the model complexity.



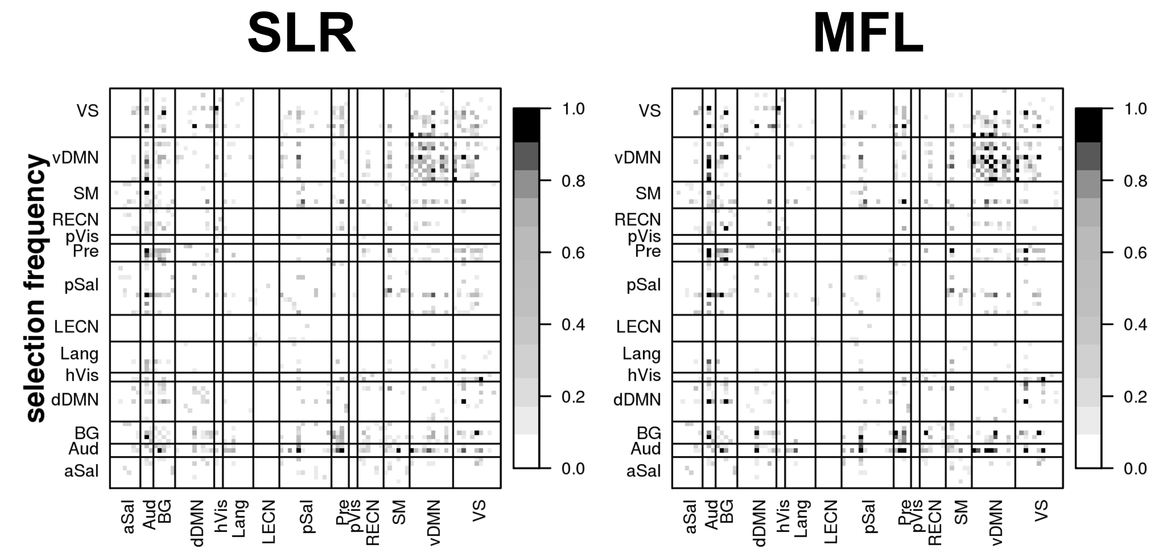
Feature stability. Comparison of MFL to reference methods: SLR, and SMLR. Fraction of features that are selected >90% in 100 bootstrap replicates. SLR and SMLR do not differ substantially. MFL has a higher fraction of stable features in models with 50-2,000 connections.



Edge weights.



Edge selection frequencies.



Conclusions

MFL provides increased stability in feature selection compared to the commonly-used sparse logistic regression models, while maintaining the same classification performance as RLR, SLR and SMRL. Most importantly, restricting the set of selected features to be the same across multiple classification tasks greatly simplifies model interpretation.

For more information contact Andre Altmann: altmann@stanford.edu