Prediction of Brain Activity Based on Sparse Regression using the Elastic Net Method

Melissa K. Carroll
Department of Computer Science, Princeton University, 35 Olden Street, Princeton, NJ 08540

Guillermo A. Cecchi, Irina Rish, Ravi Rao, and Rahul Garg
Computational Biology Center, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598
(Dated: October 15, 2007)

Note: lead author is a full-time PhD student requesting travel support.

There has recently been great interest in using predictive modeling to infer the instantaneous cognitive state of a subject, e.g. viewing an image of a face or reacting to a barking dog in a video game, using only the subject's fMRI data. Such modeling is typically achieved using machine learning approaches, in which models are trained on fMRI data for which the subject's cognitive state over time is known and has been labeled. While such approaches have demonstrated that cognitive states can be predicted using fMRI, prediction accuracy can be improved. Also, relatively little effort has been put into translating these predictive models into models of brain dynamics, sparking interest in improving the interpretability of predictive models.

In the regression paradigm, a recent technique from bioinformatics, the Elastic Net [2], attempts to address prediction and interpretation concerns using a novel regularization scheme. In addition, this method offers greater flexibility than traditional regression methods, enabling control over model properties.

Elastic Net effectively optimizes the functional:

$$L(\lambda_1, \lambda_2, \beta) = ||\mathbf{y} - \mathbf{X}\beta||_2^2 + \lambda_1 ||\beta||_1 + \lambda_2 ||\beta||_2^2, \qquad (1)$$

thus minimizing both the least-squares model error and the L_1 and L_2 norms of the model coefficients, The goal is to achieve the sparse modeling benefits of L_1 regularization while using L_2 regularization to relax exclusion of correlated predictors, leading to a more accurate model of the underlying data structure. Increasing the L_1 regularization leads to more voxels being selected; increasing the L_2 regularization leads to more inclusion of correlated voxels.

We evaluate the Elastic Net for fMRI, focusing on the effects of increasing the L_2 regularization on the sparse model properties. Data were obtained from a predictive modeling competition [1]. Noting that voxels nearby in space are frequently correlated, we evaluate the spatial distribution of a model as a proxy for the degree of exclusion of correlated voxels. Our findings suggest that the most predictive and robust models use highly distributed information while drawing adequately from relevant spatially localized clusters of correlated voxels. These results highlight the non-trivial relationship between localization and distribution in functional brain processing and exemplify the flexibility of the Elastic Net method for controlling model performance and interpretation.

[1] Pittsburgh EBC Group. PBAIC Homepage: http://www.ebc.pitt.edu/2007/competition.html.
[2] H. Zou and T. Hastie. Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society, Series B*, 67(2):301–320, 2005.

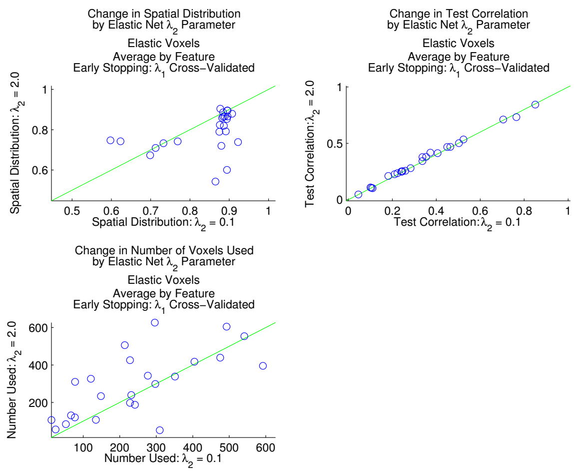

FIG. 1: Indeed, increasing the L_2 regularization decreases spatial distribution but slightly decreases sparsity, while minimally impacting prediction performance.

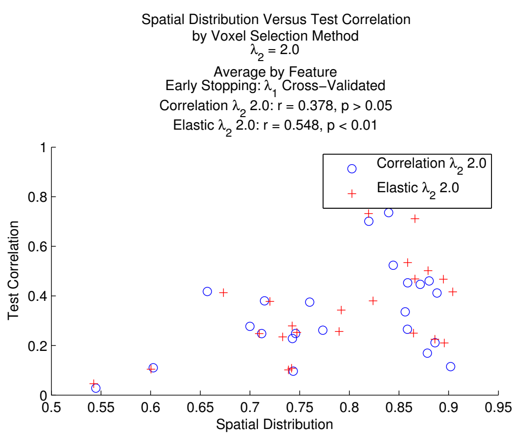

FIG. 2: If the degree of sparsity is controlled with cross-validation, mental states for which more distributed representations are found tend to be better predicted.

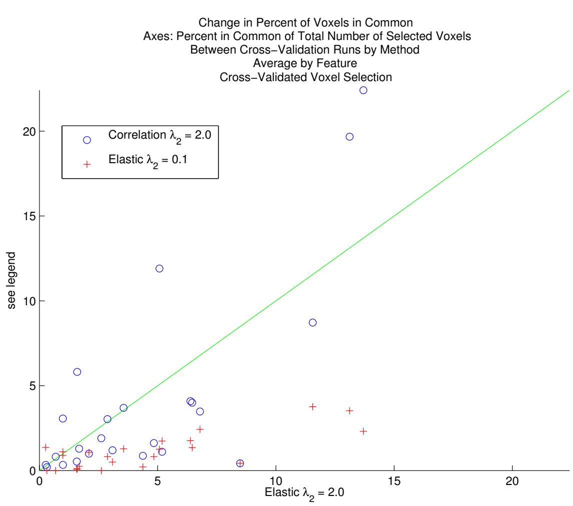

FIG. 3: Increasing the L_2 regularization results in more robust models.