ICML2016
International Conference on Machine Learning
June 19 – June 24, 2016, New York, United States

Reviews For Paper

Paper ID	828
Title	Hierarchical Region-Network Sparsity for High-Dimensional Inference in Brain Imaging

Masked Reviewer ID: Assigned_Reviewer_4

Review:

Question	
Summary of the paper (Summarize the main claims/contributions of the paper.)	The brain processes information both locally, in distinct regions of the brain, and globally as distinct regions interact. Functional magnetic resonance imagining (fMRI) is one method of recording the ongoing activity of an entire brain while a human subject is engaged in a task, and provides data with which to determine the local regions and networks important for different tasks. One method of determining which regions or networks are important for a task is to train a classifier which takes as input the values of voxels in different conditions and attempts to classify what the subject was doing or seeing at the time brain activity was recorded. By looking for regions in the brain which the classifier weights heavily, we hope to determine which areas of the brain are important for certain tasks. Typical recordings made in this way record from ~100,000 voxels while the number of training examples available is much smaller. Regularization is therefore required to fit classifiers and the authors introduce a grouped sparsity inducing prior which can allow voxels to be preassigned into regions and networks, which are then penalized in a manner similar to the work of Yuan and Lin, 2006.
Clarity (Assess the clarity of the presentation and reproducibility of the results.)	Below Average
Clarity - Justification	The paper seems to assume a fair amount of background fMRI knowledge that I suspect a typical ICML reader may not have, which renders the paper somewhat difficult to follow. For example, it was not clear to me until I was well into the paper that instead of learning which regions and networks to assign a voxel to this was knowledge that a researcher would explicitly incorporate into the prior. Additionally, for a general audience, it would be helpful to include background and motivation as to why one versus rest classification is useful for determining the regions and networks involved in different tasks. It would also be helpful to be provided with more information about the datasets used - they are described as being labeled with ongoing psychological processes, but to a general reader, it may not be clear what this means. Similarly, it was difficult to fully follow the results. It section 3.1, it is reported that across analyses the proposed method was successful at distinguishing unseen neural activity maps. It is not clear what exactly across analyses means and if this was per subject or pooled? Similarly, in section 3.2, it was not clear what out of sample was referring to? Was an out of sample data point, a point from an entire subject that was not in the training data? Additionally, the authors refer to the class mean - which presumably is a mean across subjects, but it was not clear exactly how it was obtained and why it is a

	meaningfully thing to compare too.
Significance (Does the paper contribute a major breakthrough or an incremental advance?)	Below Average
Significance - Justification	It is not clear to me that the paper contribute a major advance. It seems that the proposed methods improves accuracy for distinguishing neural activity maps but it is not clear how meaningful the small percentage difference is. Classification accuracies are roughly similar for different tradeoffs between the region vs network sparsity penalties, while the support recovered under the different settings can be markedly different. This is useful, but it is not clear how important this is in practice as you need to specify the regions and networks a priori. The concern I have is that if a region or network structure is incorrectly specified, one may be led to spurious conclusions. It would seem more impactful if the method was able to learn the groupings of voxels into regions and networks based on data.
Detailed comments. (Explain the basis for your ratings while providing constructive feedback.)	For a general ICML audience, the machine learning novelty seems relatively incremental, while it is not clear a general audience will be able to fully follow and appreciate the application of the proposed methods as they are currently presented. It would be helpful for the authors to provide more background for a general audience to follow and understand their results and place their results in context. Smaller points: The authors state that the prior establishes a tree structure. I see how one could structure the regions into networks that form a tree structure, but it seems to be that the constraints on line 248 also allow for a region to be placed into two different networks where one network is not necessarily a descendent of the other.
Overall Rating	Weak reject
Reviewer confidence	Reviewer is knowledgeable

Masked Reviewer ID: Assigned_Reviewer_5

Review:

Question	
Summary of the paper (Summarize the main claims/contributions of the paper.)	The authors present a way to incorporate brain region and network priors into the regularizer of regression/classification tasks. The evaluate their model on neuroimaging data from the human connectome project.
Clarity (Assess the clarity of the presentation and reproducibility of the results.)	Above Average
Clarity - Justification	The paper is clearly written, previous work is acknowledged, and the experiments seem reasonable.
Significance (Does the paper	

contribute a major breakthrough or an incremental advance?)	Below Average
Significance - Justification	This is a straightforward modification of a regularizer based on group sparsity regularizers. There is generally nothing against that, if the domain specific extra knowledge improves the performance of the algorithms. However, this does not seem to be the case as the method performs on par with standard Lasso or Elastic Net (Fig. 3).
Detailed comments. (Explain the basis for your ratings while providing constructive feedback.)	I think it is a good idea to use such kind of regularizers in the analysis of neuroimaging data. However, it doesn't seem to help. This is interesting by itself, but still makes it hard to argue why the paper should be accepted to ICML.
Overall Rating	Weak reject
Reviewer confidence	Reviewer is knowledgeable

Masked Reviewer ID: Assigned_Reviewer_6

Review:

Question	
Summary of the paper (Summarize the main claims/contributions of the paper.)	The paper considers high-dimensional logistic regression for task classification from neuroimaging data. A structured penalization scheme is considered that combines group/hierarchical lasso penalties capturing prior information on both brain networks and brain regions thereby bridging two principles of neurobiological architecture: functional segregation and functional integration. The method is demonstrated extensively using data from the Human Connectome Project in terms of classification accuracy, sample complexity and support recovery.
Clarity (Assess the clarity of the presentation and reproducibility of the results.)	Excellent (Easy to follow)
Clarity - Justification	The paper is very well written. The literature is nicely reviewed and the concepts are introduced in sufficient details so that they can be understood by both neuroscience and ML communities. The experiments are clearly explicated.
Significance (Does the paper contribute a major breakthrough or an incremental advance?)	Above Average
Significance - Justification	This is a very nice application paper. While the proposed method in and of itself has limited novelty as it builds upon existing group/hierarchical lasso approaches, the application to the biological problem at hand is very well developed and interesting.
Detailed comments. (Explain the basis	The paper evaluates the thesis that combining two key principles guiding brain organization: functional and regional architectures is beneficial for improving classification from neuroimaging data. This is accomplished by leveraging recent work on structured group lasso/ hierarchical lasso penalties.

for your ratings while providing constructive feedback.)	The ideas and method presented in the paper are clearly exposed, justified from a biological standpoint and evaluated. Rigorous and extensive background is provided which is valuable to both ML/neuroscientist readers which might not necessarily be familiar with either the application considered or the structured sparsity approaches, respectively. The experiments are thorough and might motivate researchers in high-dimensional estimation to develop novel methods for this problem.
Overall Rating	Weak accept
Reviewer confidence	Reviewer is knowledgeable