Sex Prediction Using Multi-task functional Connectomes

Richard Wang, Javid Dadashkarimi[1], and Dustin Scheinost[2]

[1] Department of Computer Science, Yale University, New Haven, CT
[2] Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, CT

Abstract. From neurotranmitters to macroscopic morphometry, men and women differ in their brain's structures and functions. These sex differences putatively underlie differences in behavior, disease risk, and treatment outcomes. A better characterization of brain-based sex differences may lead to better understanding of risk and, eventually, more targeted treatments. Previous research mostly focus on either explanatory analysis to find out areas of brain that are different between sexes or use resting connectomes to predict sex. In this paper first we show that each task has specific impact on sex prediction. Second, combining all task connectomes leads predictory models towards a better prediction while introducing a better generalizability. Experimental results on two standard collections including Human Connectome Project and Philadelphia Neurodevelopmental Cohort show that combined connectome achieves up to 94% accuracy.

1 Introduction

Sex differences are widely studied in neurosceince. Studies have shown that males and females display differences in brain volume, surface area, cortical thickness, white matter microstructure, and functional connectivity[9]. Because many psychiatric disorders differ substantially by sex[7], understanding sex differences can be vital to developing effective treatment against these disorders for both sexes.

Functional connectivity derived from functional magnetic resonance imaging data has been extensively used to characterize sex and generally group differences [5]. While some people used explanatory analysis to explain differences (e.g., [2]) several studies use predictory models to this aim (e.g., [5, 6**?**]). In some collections of data with hundreds of subject, sometimes it is hard to find underlying causal relationship with explanatory model due to variation across subjects, and signal to noise ratio. On the other hand, using predictory model, though not gauranteed to find robust true relationship between nodes, we can construct hypothesis and provide a pathway to find possible relationship between nodes.

Current literature uses resting scan data as input for prediction while people report over fitting becausoe of limited data [6]. We propose to utilize functional

data, combine them with resting state to create a greater input for the predictory model rather than exploring new participants or presumably new sites for scan. We seek to find out if using multiple tasks will increase the sex prediction accuracy.

Experimental results on two standard datasets including Human Connectome Project and Philadelphia Neurodevelopmental Cohort show that task connectomes contribute differently on sex prediction while combining them with resting data consistently achieves reasonably high level of accuracy up to 94%.

2 Predictory Models, Individual and Combined functional Connectomes for Predicting Sex

In this section we introduce a couple of predictory models widely been used in functinal connectomic literature and support vector machines (SVM), a well known supervised model in machine learning, and move towards a hypothesis of using either resting data, task connectomics, or combining them to achieve a better prediction.

Predictory Models Connectome based predictory models (CPM) based on fMRI data has extensively been used to predict participants' traits [8]. CPM taking advantage of two main steps including one feature selection and one model fitting. In the first step, features that are positively correlated with labels and features with negative correlation are identified from a number of training data. Generally speaking, each feature is defined by edge intensity of brain that is activated during the scan. In the second step CPM fits a linear model, one intercept and once slope, explaining relationship between features and labels. Finally, this model applies on testing data by masking over the positive/negative features and feeding into the linear model. Quantitative differences between predicted values and actual ones exhibit the level of performance your model achieved. While promising, people suggest to use some regularization terms in the model to achieve a better generalizability [4, 3, 10].

On the other hand, SVM has been extensively used as a supervised machine learning model, that divides the data points into different classes [1]. To this end, SVM considers several hyperplanes and selects ones with biggest margin between closest data points: SVM constructs a hyperplane, and createse a linear model to predict output. SVM chooses the closest supporter vectors that have the maximized distance, making the hyperplane fare to both classes:

$$\hat{w} = \arg\min_{w} ||w|| \tag{1}$$
$$\text{subject to } y_i(wx_i - b) \geq 1, \text{for all } 1 \leq i \leq n$$

where w is weights, x_i is feature, and y_i is the appropriate label.

Combining Connectomes Regardless of which model we use for prediction, choosing a feature space is a crucial step. Previous works are mostly rely on using resting connectomes where we suggest to use task-based connectomes and

ID	Collection	#male	#female	size	age	#tasks
HCP	Human Connectome Project	241	274	515	28 ± 3.98	9
PNC	Philadelphia Neurodevelopmental Cohort	251	320	571	15 ± 3.65	3

Table 1: Characteristics for the HCP and PNC datasets.

ultimately a combined feature space. Roughly speaking, lets say we have N participants and each one has feature dimensionality of $\mathbb{R}^m$ where each element is edge intensity of of a connection in the brain. We can use either resting data (i.e., $X \in \mathbb{R}^{N \times m}$) or task data with similar dimensionality. Here we also suggest to use a combination of all tasks and rests (i.e., $X \in \mathbb{R}^{N \times km}$) where k is the total number of tasks and rests.

3 Experiments

3.1 Experimental Settings

We take in a 4 dimensional correlation matrix $\mathbb{R}^{268 \times 268 \times 515 \times 9}$ is the correlation matrix, 515 is the amount of subject and 9 is the amount of tasks.We first collapsed the correlation matrix, cut it in half and create a vector the size of $\frac{268 \times 267}{2} = 107334$, getting rid of repeating elements and correlations between node itself. Using a 10-fold cross validation method, we divide the data set into training and testing of ratio $9:1$. The chosen predictory model first correlates the edges with phenotypical measure(here it is sex difference); we then build a model base on selected significant edges from the training set with P value less than 0.05. Last, we run the model on test data set, and create a one dimensional matrix of value between 1 and 0 representing the subjects' predicted sex(greater than or equal to 0.5 represent male, less than 0.5 represent female). We repeated the steps m+1 times, with m being the amount of functional task the data set has. The result in table 2 and table 3 represent the result of the prediction of HCP and PNC, independently.

To test out the effect of combining function tasks, we conducted 27 different iterations, of 2 different function tasks combined, and we recorded the top 10 predictions (seen in table). We then change the method of selecting edges; instead of choosing edges with a threshhold(P value) higher than 0.05, we repectively choose the top five percent, ten percent, and fifteen percent edges to build the model.

we investigate each task's contribution to the overall prediction. We again set the edge selection threshold to 0.05, and calculate what percentage of edges each task contribute.

Table 1 shows the dataset descriptions we used in our experiments.

3.2 Experimental Results

As seen in Table 1 and 2, the individual task prediction for all three models significantly predict sex differences within data set.(pvalue, to be calculated) SVM model outperforms CPM and rCPM on single task predictions. rCPM is slightly worse than SVM, but generic CPM model has signifigantly lower prediction compares to the other two models. The combined task predictions significantly outperform single task prediction, with SVM and rCPM predictions going up to 94 percent. This can be applied to both datasets,However, there's an exception in PNC dataset, whereas the combined task prediction for CPM did not improve from single task prediction. The nback task even yield a higher prediction rate than combined task prediction.

duo task performance To further study task contribution to model prediction, we see a significant increase of prediction accuracy when 2 tasks are combined. the prediction rate increased significantly, SVM prediction accuracy for duo tasks are all above 90 percent. There's signifigant correlation between combining function tasks and the prediction accruacies.

3.3 edge selection

Different tasks contribute different amount of edges toward combined task prediction. Tasks that contribute the most to the prediction, tend to be the ones that yield higher individual task predictions. As seen in Figure 1, rest 1 and rest 2 overwhelming contribute more edges toward the final prediction. This is most obvious in SVM model, with rest 2 yielding highest prediction rate amoung all individual tasks. (tbd)

3.4 generalization of models

Table 4(change later) shows generalization of CPM, SVM, and rCPM to corss data set prediction, using different percentage of edges(figure 1). The CPM model result is very unpleasant, having prediction result within 2 percent of null prediction. However SVM and rCPM both produced predictions around or above 70 percents, showing model tolerance to over fitting problem. Interestingly, SVM and rCPM both have predictions in the same 69-77 percent range, and both of them did better with PNC being the training set, and hcp being the testing set. CPM model displays the exact opposite.

All three models experience an increase in prediction accuracy, but only when we trains HCP to predict PNC. When it's the other way around, there's absolutely no change in prediction accuracy with regard to edge selections. It's hard to conclude correlation between edge contribution and prediction accuracy.

	gam	rest 1	rest 2	lang	motor	relation	social	wm	emo	all
cpm	72.816	63.301	60.388	69.515	69.320	68.738	**75.534**	70.097	69.709	**81.165**
svm	84.854	87.767	**91.068**	86.990	88.932	87.573	86.602	85.049	85.049	**94.757**
rcpm	**87.379**	78.252	80.194	76.700	68.932	82.913	82.330	78.252	71.456	**94.757**

Table 2: HCP dataset precision results. HCP is consisted of 268 nodes, with We rounded the predictions by CPM and rCPM into 0 and 1.

	emotion	nback	rest	all
cpm	64.45	**67.43**	62.87	65.85
rcpm	**79.34**	74.61	75.83	84.41
svm	81.61	**82.14**	76.36	85.11

Table 3: PNC dataset precision results. PNC consisting of 250 nodes rather than 268 and is consisting of subjects in in range of age 15 $\pm$ 3.65.

	train	test	svm	cpm	rcpm
5%	hcp	pnc	71.45	57.79	69.35
	pnc	hcp	76.89	54.76	75.53
10%	hcp	pnc	72.33	59.02	71.98
	pnc	hcp	**76.89**	53.98	77.48
15%	hcp	pnc	73.20	59.71	73.03
	pnc	hcp	76.89	53.79	77.48

Table 4: cross-dataset results.

	DH	DJ	JK	DI	HJ	DG	HK	GJ	DK	DL
cpm	78.641	77.670	77.476	76.893	76.699	76.117	75.922	75.728	75.340	74.951

Table 5: Duo functional tasks prediction result. Using HCP dataset, CPM model.D=gam, E=rest1, F=rest2, G=lang, H=motor, I=relation, J=social, K=wm, L=emo.

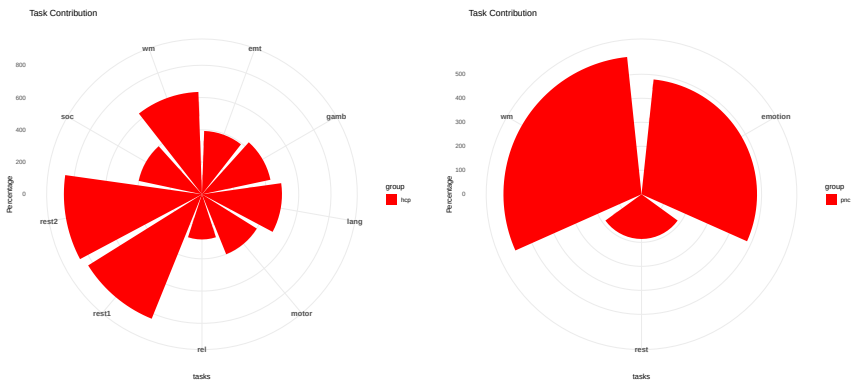
	EH	FH	HI	FI	HK(TS)	HJ(TS)	FJ	EJ	EF	DH
SVM	93.981	93.786	93.010	92.816	92.427	92.427	91.845	91.650	91.456	91.262

Table 6: Duo functional tasks prediction result. Using HCP dataset, SVM model.

4 Conclusion

In this work, we proposed combining function tasks to improve functional tasks connectomes. We tested how combining 2 tasks will improve prediction accuracies. We used three seperate model, CPM, SVM, and rCPM, to predict sex of our subjects. We predicted sex using individual functional tasks, as well as duo functional tasks, and all-combined tasks.

Looking at task contributions, we found that different tasks contribute to the model differently. Very interestingly, in HCP data set resting task 1 and 2 con-



(a) hcp (b) pnc

Fig. 1: Each functional task edge contribution to overall prediction.

tributed the most to the model, while in PNC data the resting task contributed the least to the model. This illustrates that for different groups, different functional tasks may have different contributions to predicting sex differences. In an effort to test our model's ability for generalization, we did cross dataset prediction, and found similar trend among SVM and rCPM model. CPM prediction fell to null prediction, hence its result is discounted. What's interesting about cross data set prediction is we see a gradual increase in prediction accuracy with accordance to percent edges selected, when HCP is the training data and PNC is the testing set. When the PNC is used as the training set, the prediction accuracy is unhinged by the amount of edges selected. This is true for all three model. A possible explanation for this is that PNC data is collected from younger participants, and sex differences are not every obvious before adolescent, thus an increase in data input does not affect the prediction accuracies.

Our study has its limitation. While in total we have 700 hundred subjects, the cross dataset prediction significantly decreased compared to within sample prediction. The CPM model prediction dropped all the way down to null prediction. The PNC dataset is consisted of people around the age of 15. Adolescent subjects have much more subtle sex differences, which could explain some over fitting in the cross dataset predictions. The model still need to be tested on larger public dataset, to ensure its ability to generalize. In summary, we presented how combining task affects the prediction accuracy in CPM, SVM, and rCPM model. Our result suggest that combination of functional tasks significantly improves

the predictions. The additional tasks contribute useful information for the model to better model after the data.

Bibliography

[1] Alex M Andrew. An introduction to support vector machines and other kernel-based learning methods by nello christianini and john shawe-taylor, cambridge university press, cambridge, 2000, xiii+ 189 pp., isbn 0-521-78019-5 (hbk,£ 27.50). *Robotica*, 18(6):687–689, 2000.

[2] Javid Dadashkarimi, Siyuan Gao, Erin Yeagle, Stephanie Noble, and Dustin Scheinost. A mass multivariate edge-wise approach for combining multiple connectomes to improve the detection of group differences. In *International Workshop on Connectomics in Neuroimaging*, pages 64–73. Springer, 2019.

[3] Siyuan Gao, Abigail S Greene, R Todd Constable, and Dustin Scheinost. Combining multiple connectomes via canonical correlation analysis improves predictive models. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 349–356. Springer, 2018.

[4] Siyuan Gao, Abigail S Greene, R Todd Constable, and Dustin Scheinost. Task integration for connectome-based prediction via canonical correlation analysis. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 87–91. IEEE, 2018.

[5] Abigail S Greene, Siyuan Gao, Dustin Scheinost, and R Todd Constable. Task-induced brain state manipulation improves prediction of individual traits. *Nature communications*, 9(1):2807, 2018.

[6] Wolfgang Huf, Klaudius Kalcher, Roland N Boubela, Georg Rath, Andreas Vecsei, Peter Filzmoser, and Ewald Moser. On the generalizability of resting-state fmri machine learning classifiers. *Frontiers in human neuroscience*, 8:502, 2014.

[7] Moffitt Rutter M, Caspi A. Using sex differences in psychopathology to study causal mechanisms: unifying issues and research strategies. *Child Psychol Psychiatry*, 44:1092–1115, 2003.

[8] Xilin Shen, Emily S Finn, Dustin Scheinost, Monica D Rosenberg, Marvin M Chun, Xenophon Papademetris, and R Todd Constable. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *nature protocols*, 12(3):506, 2017.

[9] Xueyi Shen Michael V Lombardo Lianne M Reus Clara Alloza Mathew A Harris Helen L Alderson Stuart Hunter Emma Neilson David C M Liewald Bonnie Auyeung Heather C Whalley Stephen M Lawrie Catharine R Gale Mark E Bastin Andrew M McIntosh Ian J Deary Stuart J Ritchie, Simon R Cox. Sex differences in the adult human brain: Evidence from 5216 uk biobank participants. *Cerebral Cortex*, 28(8):2959–2975, 2018.

[10] Larry Wasserman. *All of nonparametric statistics*. Springer Science & Business Media, 2006.