Predicting Mental Functions from Brain Activations

Madhura Parikh[*]
Department of Computer Science
University of Texas, Austin
mparikh@cs.utexas.edu

Subhashini Venugopalan[*]
Department of Computer Science
University of Texas, Austin
vsub@cs.utexas.edu

Abstract

Over the past few decades neuroscientists have studied brain images from EEG/MEG, fMRI and other sources to identify associations between psychological tasks and activity in brain regions [**?**]. Although these studies have led to large amounts of literature and several discoveries of cognitive functions associated with certain brain regions (or networks) the mapping between functions to brain regions and vice-versa still remains largely unclear. For the purposes of this project, we look at enhancing a new automated framework NeuroSynth [**?**] that combines text-mining and machine learning techniques to generate probabilistic mappings between cognitive and neural states. Starting from their Naive Bayes classifier, we apply more sophisticated binary classifiers to the problem and also consider multi-label predictions and transfer-learning(?).

1 Introduction

1.1 Our Contribution

2 Related Work

3 Methods

Notation : We denote matrices by boldface capital letters $\mathbf{X}$ and vectors by bold-face lower case letters $\mathbf{x}$. The set of real valued D dimensional vectors are denoted by $\mathbb{R}^D$. The set of (target) labels are denoted by the letter $\mathcal{L}$, and $|\mathcal{L}|$ denotes the cardinality of the set of labels.

Our input data consists of brain volume images. We denote by $\mathbf{x}_n$ the nth brain volume with voxels collected in a real valued D dimensional vector. The total number of brain volumes is represented by N. Each brain volume is associated with a set of labels $\mathcal{L}_n = l_1, \ldots, l_K$ chosen from the full set of possible target labels $\mathcal{T} = \cup_{n=1,\ldots,N}\mathcal{L}_n$ with $|\mathcal{L}| = L$.

We approach the task of predicting brain functions as a classification task and address it using the following techniques:

3.1 Single Label Prediction

As a first step to measure the feasibility of the task we consider the task of predicting one of two labels for each brain volume. The single label prediction learns a mapping $f : \mathbf{x}_n \to \mathcal{L}_n$, such that $|\mathcal{L}_n| = 1$ for all n, where $|\mathcal{T}| = |\cup \mathcal{L}_n| = 2$. This essentially learns a simple binary classifier for all pairs of labels. Each classifier is trained and tested on an appropriate subset of the data where the labels for the brain volume are unique. We experimented with the following binary classifiers:

---

[*]Authors contributed equally

Naive-Bayes We built pairwise naive-bayes classifiers as a baseline to replicate the results presented in [**?**].

Logistic Regression An l_1 regularized logistic regression was our next model of choice for a more powerful classifier.

Linear SVM We use a linear SVM which scales well with the size of our features. The hyperparameter C is optimized by a grid search in the space $C = \{1, 10, 100, 1000\}$.

Ensemble We finally build an ensemble of the above three classifiers and consider the label that is the majority in the output of the above three classifiers.
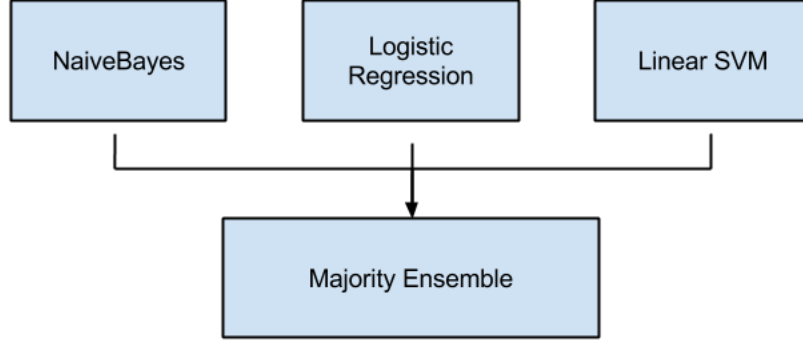

Figure 1: We use naive-bayes, logistic regression, SVMs and a majority ensemble for pairwise binary classification of the labels.

3.2 Multi-Label Prediction

Our next approach is to perform multilabel classification to predict multiple functions for a given brain volume. This learns a mapping $f : \mathbf{x}_n \to \mathcal{L}_n$, over the full set of labels (where $|\mathcal{L}_n| \geq 1$) and $|\mathcal{T}| = |\cup \mathcal{L}_n|$. [**?**] suggests several methods to perform multilabel classification which include label projection, label ranking and label decomposition. We prefer the label decomposition method for it's ease of interpretation. In particular we choose the *One-vs-All* decomposition approach. This method learns multiple binary classifiers ($|\mathcal{T}|$ to be precise), one for each label to predict the presence/absence of that label. Our experiments used the following binary classifiers for the multi-label prediction task:

Naive-Bayes We built pairwise naive-bayes classifiers as a baseline to replicate the results presented in [**?**].

Logistic Regression We build two models here. The first with l_1 or *lasso* regularization and the other with l_2 or *ridge* regularization.

Linear SVM In a manner similar to our single label prediction, we use a linear SVM by setting the hyperparameter C via a grid search in the space $C = \{1, 10, 100, 1000\}$.

Ward Clustering Since the number of features in our data is extremely large, we do a dimensionality reduction of the features by first performing ward clustering [citation] and then mapping the points to different clusters. The number of clusters is 5000. We then apply a logistic regression classifier with l_1 regularization on the reduced data.

Unique train Since our data is quite sparse and there's a high possibility of noise in data with multiple labels, we build what we call a unique train model based on logistic regression. In the unique train model, we identify all brain volumes with a single unique label, and train a classifier for each label using only this data. During test, we use the One-vs-All approach to test on multilabel brain volumes.

All ones Another baseline we consider is the all ones baseline, which predicts ones for all labels on all input data points. Such a manipulation will give result in 100% recall for the multilabel task, but very low precision values.

In addition to the above, we did experiment with PCA and k-nearest neighbours for $k \in \{2, 3, 4\}$ but the models were not trained similarly or performed extremely poorly and hence we do not report results for these.

2

Multi-label Classification and Transfer Learning
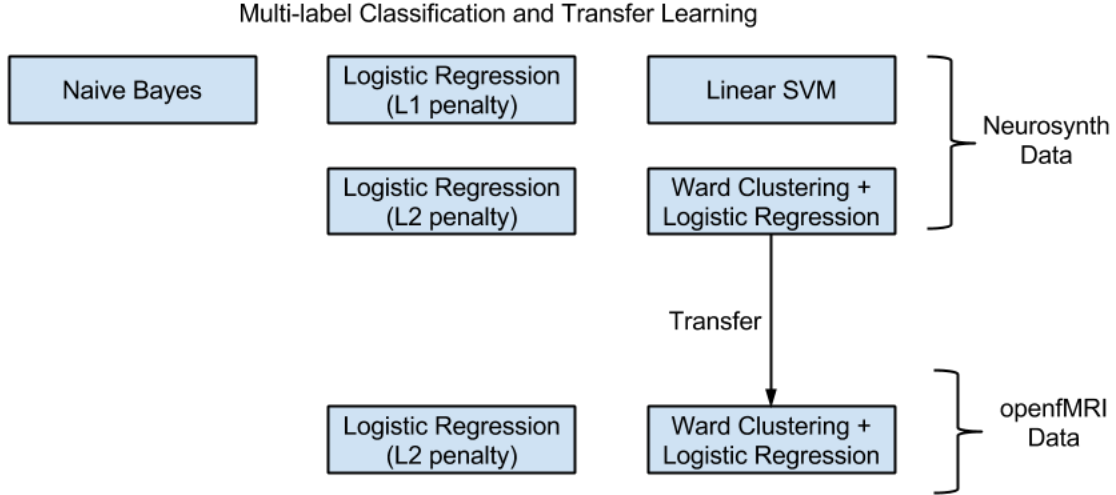
Figure 2: We experiment several classifiers for the multilabel task. We pick our best performing models, namely logistic regression with l_2 penalty and the ward clustering based models and evaluate their performance on the openfMRI data.

3.3 Transfer Learning

As mentioned in Section **??**, one of the main purposes of this particular synthesis method for predicting brain functions is to be able to transfer the models from the synthesis domain to the real openfMRI domain. Based on the mapping functions that we learned in Section 3.2 we want to apply the same function or a small modification of the function to new studies. That is, apply the mapping $f : \mathbf{x}_n \rightarrow \mathcal{L}_n$, to $\mathbf{x}'_n$ where $(\mathbf{x}'_n \notin \mathbf{X})$. In fact, our features $\mathbf{x}'_n$ are different both in dimension and constitution (the elements are continuous real valued data unlike the binary synthesis data). More details about the pre-processing is mentioned in Section **??**. In addition, we are interested in a predicting a different set of labels or functions, as constrained by the availability of data in the new domain. In effect, the transfer learning is attempting to learn $f' : \mathbf{x}'_n \rightarrow \mathcal{L}_n$. We experiment with the following methods here:

Ward Clustering We reuse the ward clustering model that we learn previously by first performing ward clustering [citation] on the data in the synthesized domain to cluster and map the points to different clusters. We then train an l_2 regularized logistic regression *One-vs-All* classifier on the data. Next we reapply the same clustering transformation to the openfMRI data and use the logistic classifier to predict the labels. The number of clusters is 5000.

Logistic Regression For the sake of comparison we train an l_2 regularized logistic regression model directly on the openfMRI data.

4 Data

4.1 Preprocessing

5 Experimental Results

Our experiments for single and multilabel prediction tasks were performed on data available in the neurosynth [1] tool. After preprocessing and filtering our inpput feature matrix $\mathbf{X}$ is of dimension $5067 \times 228,453$, the target label matrix $\mathbf{Y}$ is of dimension 5067×23. The transfer learning component is tested on data from the openfMRI [2] project with feature dimensions $\mathbf{X}'$ ($479 \times 174,264$) and targer label matrix $\mathbf{Y}'$ of dimensions 479×19.

---

[1] www.neurosynth.org
[2] https://openfmri.org

3

Data samples size (4000). We use 10 fold cross validation to evaluate all our models unless otherwise indicated. We evaluate the performance of our models on the metrics such as *accuracy, precision, recall, hamming loss, F1 score* which are frequently applied to this task.

$$\text{Accuracy} =$$

5.1 Single Label Prediction

We present the accuracies of the pairwise single label predictions using all our approaches in the heat maps below.

(b) Plotting Correlations

5.2 Multi-Label Prediction

In addition to reporting the accuracy, we report precision, recall, F1 score and hamming loss. And one correct

$$\text{Precision} = \qquad\qquad\qquad\qquad \text{Recall} =$$

$$\text{F1 score} = \qquad\qquad\qquad\qquad \text{Hamming loss} =$$

$$\text{Top correct} = \frac{1}{N} \sum_{n=1}^{N} \mathbb{I}(\mathcal{Z}_{n,max} \in \mathcal{L}_n)$$

$\mathbb{I}$ denotes the indicator function and $\mathcal{Z}_{n,max}$ denotes the label predicted with highest probability for the brain volume n. The top correct value is the fraction of the data points on which the model's best predicted label (with maximum probability) was present in the true label set.

This table reports the results of the multilabel classification for the entire label set. Note here that there are 2^{22} possible label sets and accuracy is a fairly hard metric to achieve

Model	Accuracy	Precision	Recall	F1 score	Hamming loss	Top correct
Naive Bayes	1.3	**22.36**	**46.04**	**26.63**	0.207	23.7
Logistic L1	3.2	18.92	14.16	14.53	0.111	27.9
Logistic L2	**4.1**	19.58	14.73	15.22	0.108	**30.0**
Linear SVM	3.7	20.34	16.40	16.24	0.115	22.5
Ward Cluster	3.4	17.53	12.99	13.43	**0.107**	29.8
Unique Train*	0.8	**33.98**	22.09	24.83	0.162	**36.0**
All Ones*	0.0	9.78	**100.0**	17.40	0.902	5.38

Table 1: Multilabel classification full tuple results. All values are reported in percentage (%) except hamming loss. Again, excepting for hamming loss where lower loss is better, for all other metrics higher values are better. *Unique Train and All Ones do not use cross validation.

5.3 Transfer Learning

6 Discussion

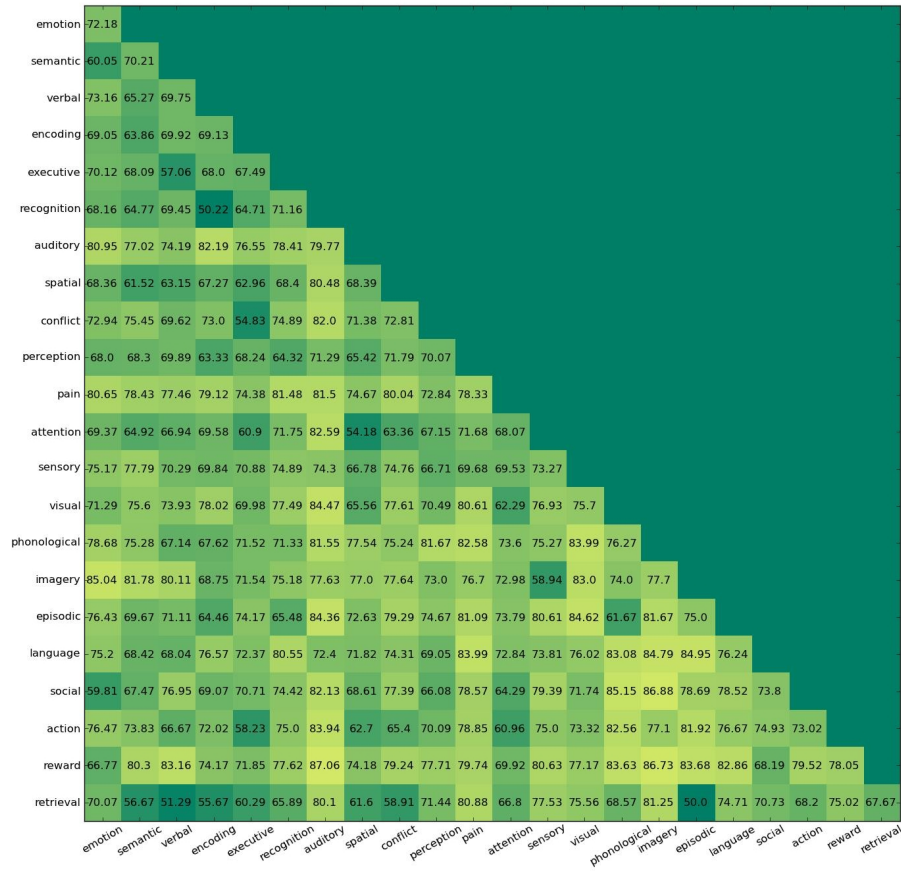– Why bad results? Or good results?

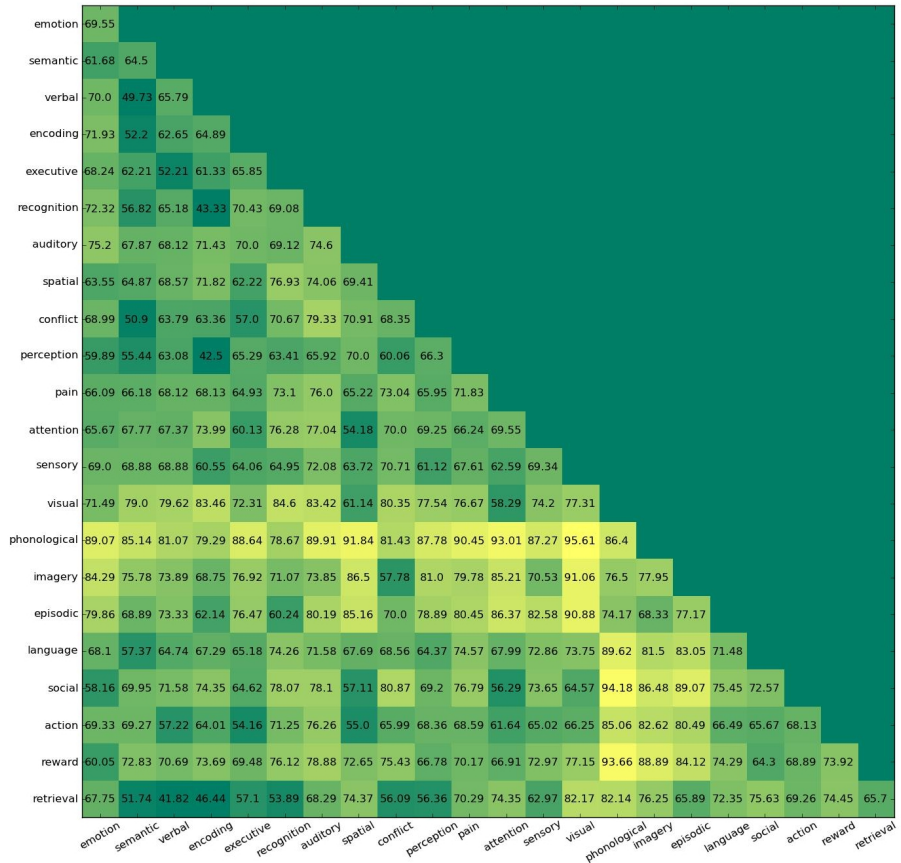Transfer - remove those studies where none of the terms are present

7 Conclusion

Acknowledgments

References

(a) Naive Bayes

	emotion	semantic	verbal	encoding	executive	recognition	auditory	spatial	conflict	perception	pain	attention	sensory	visual	phonological	imagery	episodic	language	social	action	reward	retrieval
emotion	72.18																					
semantic	60.05	70.21																				
verbal	73.16	65.27	69.75																			
encoding	69.05	63.86	69.92	69.13																		
executive	70.12	68.09	57.06	68.0	67.49																	
recognition	68.16	64.77	69.45	50.22	64.71	71.16																
auditory	80.95	77.02	74.19	82.19	76.55	78.41	79.77															
spatial	68.36	61.52	63.15	67.27	62.96	68.4	80.48	68.39														
conflict	72.94	75.45	69.62	73.0	54.83	74.89	82.0	71.38	72.81													
perception	68.0	68.3	69.89	63.33	68.24	64.32	71.29	65.42	71.79	70.07												
pain	80.65	78.43	77.46	79.12	74.38	81.48	81.5	74.67	80.04	72.84	78.33											
attention	69.37	64.92	66.94	69.58	60.9	71.75	82.59	54.18	63.36	67.15	71.68	68.07										
sensory	75.17	77.79	70.29	69.84	70.88	74.89	74.3	66.78	74.76	66.71	69.68	69.53	73.27									
visual	71.29	75.6	73.93	78.02	69.98	77.49	84.47	65.56	77.61	70.49	80.61	62.29	76.93	75.7								
phonological	78.68	75.28	67.14	67.62	71.52	71.33	81.55	77.54	75.24	81.67	82.58	73.6	75.27	83.99	76.27							
imagery	85.04	81.78	80.11	68.75	71.54	75.18	77.63	77.0	77.64	73.0	76.7	72.98	58.94	83.0	74.0	77.7						
episodic	76.43	69.67	71.11	64.46	74.17	65.48	84.36	72.63	79.29	74.67	81.09	73.79	80.61	84.62	61.67	81.67	75.0					
language	75.2	68.42	68.04	76.57	72.37	80.55	72.4	71.82	74.31	69.05	83.99	72.84	73.81	76.02	83.08	84.79	84.95	76.24				
social	59.81	67.47	76.95	69.07	70.71	74.42	82.13	68.61	77.39	66.08	78.57	64.29	79.39	71.74	85.15	86.88	78.69	78.52	73.8			
action	76.47	73.83	66.67	72.02	58.23	75.0	83.94	62.7	65.4	70.09	78.85	60.96	75.0	73.32	82.56	77.1	81.92	76.67	74.93	73.02		
reward	66.77	80.3	83.16	74.17	71.85	77.62	87.06	74.18	79.24	77.71	79.74	69.92	80.63	77.17	83.63	86.73	83.68	82.86	68.19	79.52	78.05	
retrieval	70.07	56.67	51.29	55.67	60.29	65.89	80.1	61.6	58.91	71.44	80.88	66.8	77.53	75.56	68.57	81.25	50.0	74.71	70.73	68.2	75.02	67.67

(a) Naive Bayes

(b) Logistic Regression

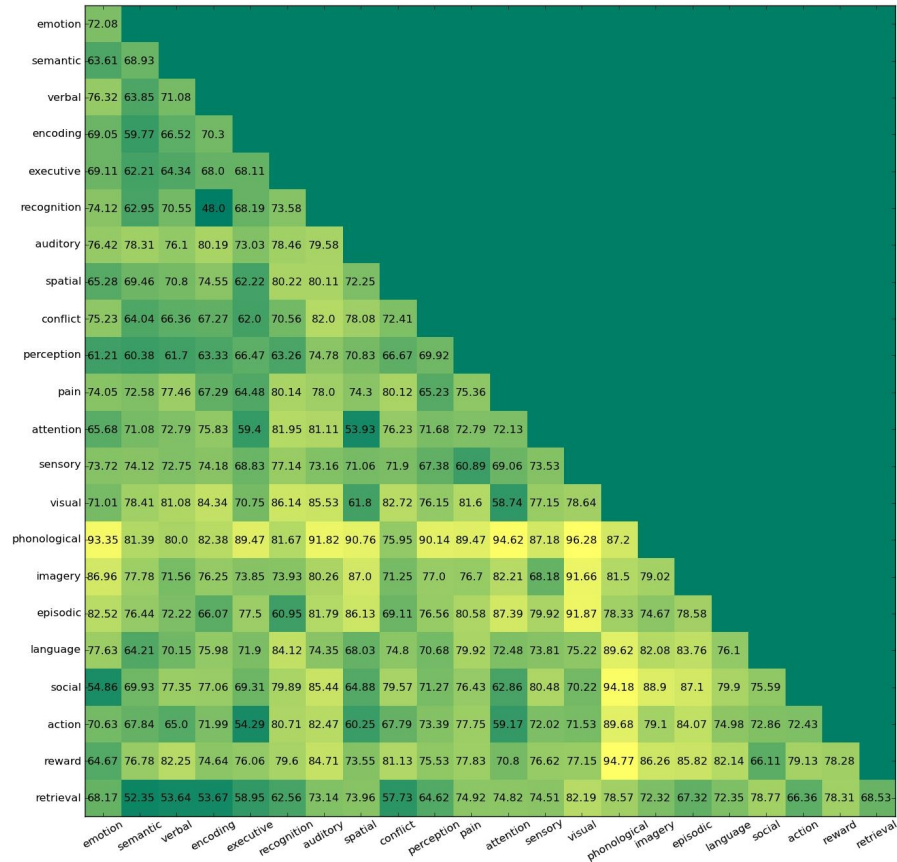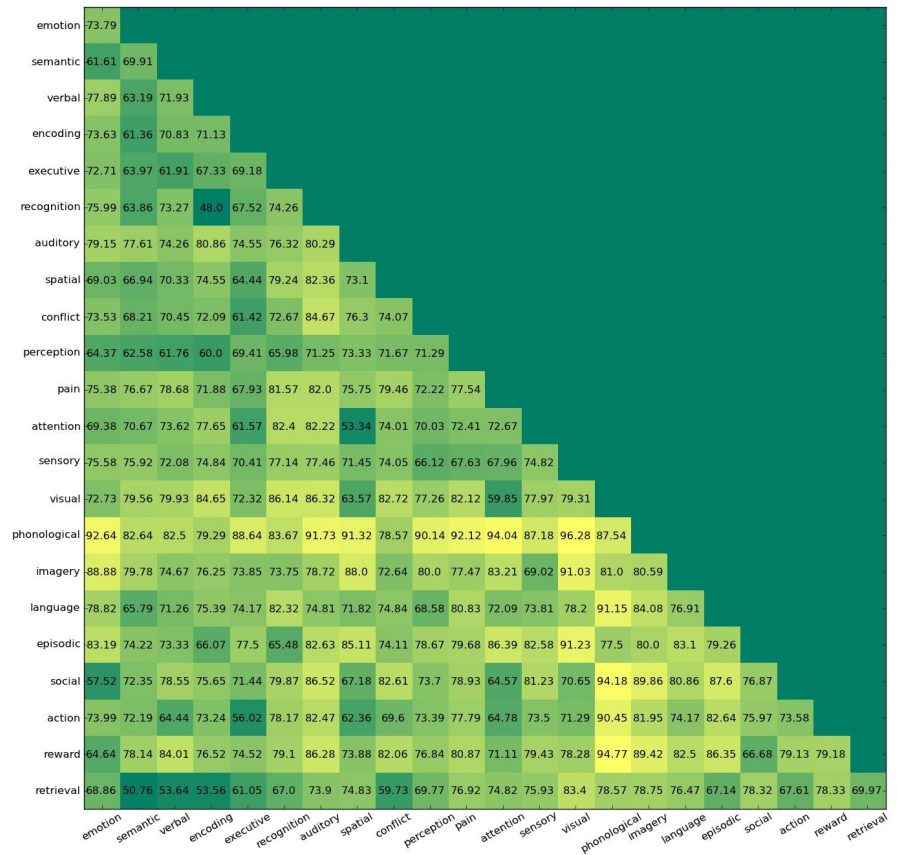|	emotion	semantic	verbal	encoding	executive	recognition	auditory	spatial	conflict	perception	pain	attention	sensory	visual	phonological	imagery	episodic	language	social	action	reward	retrieval
emotion	69.55																					
semantic	61.68	64.5																				
verbal	70.0	49.73	65.79																			
encoding	71.93	52.2	62.65	64.89																		
executive	68.24	62.21	52.21	61.33	65.85																	
recognition	72.32	56.82	65.18	43.33	70.43	69.08																
auditory	75.2	67.87	68.12	71.43	70.0	69.12	74.6															
spatial	63.55	64.87	68.57	71.82	62.22	76.93	74.06	69.41														
conflict	68.99	50.9	63.79	63.36	57.0	70.67	79.33	70.91	68.35													
perception	59.89	55.44	63.08	42.5	65.29	63.41	65.92	70.0	60.06	66.3												
pain	66.09	66.18	68.12	68.13	64.93	73.1	76.0	65.22	73.04	65.95	71.83											
attention	65.67	67.77	67.37	73.99	60.13	76.28	77.04	54.18	70.0	69.25	66.24	69.55										
sensory	69.0	68.88	68.88	60.55	64.06	64.95	72.08	63.72	70.71	61.12	67.61	62.59	69.34									
visual	71.49	79.0	79.62	83.46	72.31	84.6	83.42	61.14	80.35	77.54	76.67	58.29	74.2	77.31								
phonological	89.07	85.14	81.07	79.29	88.64	78.67	89.91	91.84	81.43	87.78	90.45	93.01	87.27	95.61	86.4							
imagery	84.29	75.78	73.89	68.75	76.92	71.07	73.85	86.5	57.78	81.0	79.78	85.21	70.53	91.06	76.5	77.95						
episodic	79.86	68.89	73.33	62.14	76.47	60.24	80.19	85.16	70.0	78.89	80.45	86.37	82.58	90.88	74.17	68.33	77.17					
language	68.1	57.37	64.74	67.29	65.18	74.26	71.58	67.69	68.56	64.37	74.57	67.99	72.86	73.75	89.62	81.5	83.05	71.48				
social	58.16	69.95	71.58	74.35	64.62	78.07	78.1	57.11	80.87	69.2	76.79	56.29	73.65	64.57	94.18	86.48	89.07	75.45	72.57			
action	69.33	69.27	57.22	64.01	54.16	71.25	76.26	55.0	65.99	68.36	68.59	61.64	65.02	66.25	85.06	82.62	80.49	66.49	65.67	68.13		
reward	60.05	72.83	70.69	73.69	69.48	76.12	78.88	72.65	75.43	66.78	70.17	66.91	72.97	77.15	93.66	88.89	84.12	74.29	64.3	68.89	73.92	
retrieval	67.75	51.74	41.82	46.44	57.1	53.89	68.29	74.37	56.09	56.36	70.29	74.35	62.97	82.17	82.14	76.25	65.89	72.35	75.63	69.26	74.45	65.7

(b) Logistic Regression

Figure 3: Pairwise binary classification accuracy for single label prediction

(a) Linear SVM

(b) Ensemble of naive bayes，logistic regression and SVM

Figure 4: Pairwise binary classification accuracy for single label prediction