Domain Adaptation with Application for Brain Decoding in Task fMRI

Shuo Zhou

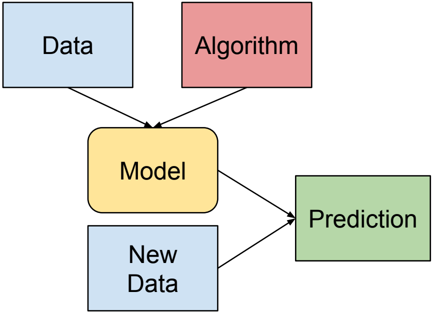
PhD candidate @ Department of Computer Science
The University of Sheffield
Regent Court, 211 Portobello, Sheffield, UK

szhou20@sheffield.ac.uk

December 23, 2020 @BNU

Brief Intro to Machine Learning

Making predictions for unlabelled (unseen) new data



Example: Spam email, credit card application ...

$$y = f(x) = ax + b$$

Theoretical Support

VC Generalisation Bound [1] for a statistical model f:

$$\epsilon_{\text{test}}(f) \leq \epsilon_{\text{train}}(f) + \sqrt{\frac{\text{Complexity}}{N}}, \text{(Eq. 1/3)}$$

- ϵ_{test}: Probability of making mistakes on unseen test data.
- ϵ_{train}: Probability of making mistakes on training data.

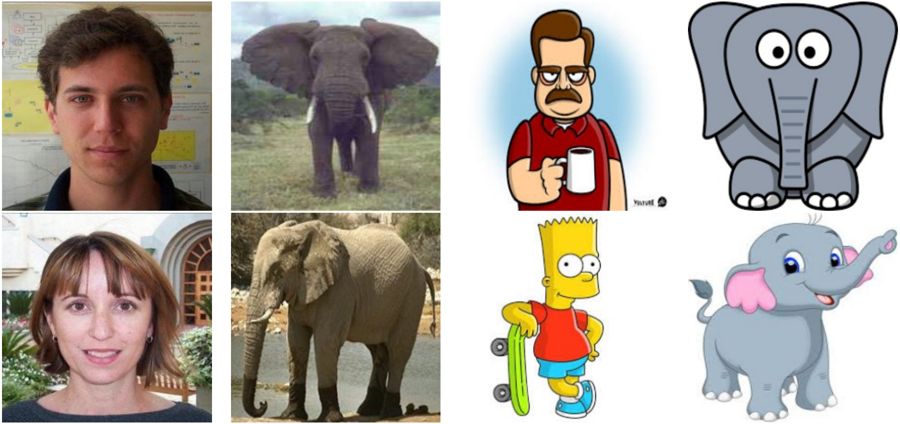
Example: Ridge regression

$$\arg\min_{\mathbf{w}} \sum_{i=1}^{n} (y_i - \mathbf{w}^\top \mathbf{x}_i) + \|\mathbf{w}\|_2^2$$

Assumption:
Training and testing samples are drawn from the same distribution $\mathcal{D}$.

[1] Vapnik, V. N.; Chervonenkis, A. Ya. 1971. On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. *Theory of Probability & Its Applications*. 16 (2).

Learning from Another Distribution?



Images from PACS dataset [1].

[1] Li, D.; Yang, Y.; Song, Y. Z.; & Hospedales, T. M. 2017. Deeper, broader and artier domain generalization. In *ICCV*, 5542-5550.

Visualisation

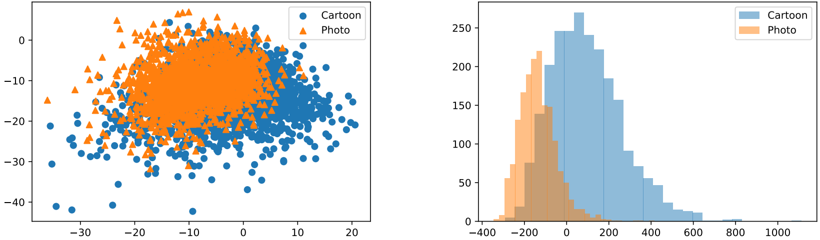


Figure: 2D/1D visualisation of Photo Cartoon images feature (extracted by AlexNet) distributions in PACS dataset via PCA

Domain Adaptation

Definitions:

- Source: Labelled domain (dataset) for training
- Target: Unlabelled domain (dataset) to make prediction

Domain adaptation is a **homogeneous** transfer learning problem:

- Same feature space
- Same learning task
- Different distributions across domains

Reference:
[1] Pan, S. J.; & Yang, Q. 2009. A Survey on Transfer Learning. *TKDE*, 22(10): 1345-1359.
[2] Redko, I. et al., 2020. A Survey on Domain Adaptation Theory: Learning Bounds and Theoretical Guarantees. *arXiv preprint* arXiv:2004.11829.

Theoretical Support for Domain Adaptation

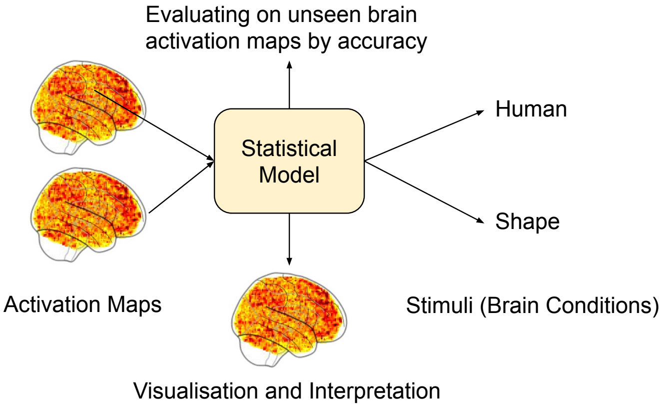
Generalisation bound [1]:

$$\epsilon_t(f) \leq \epsilon_s(f) + \underbrace{\mathrm{Dist}(\mathcal{D}_t, \mathcal{D}_s)}_{\text{Distribution Divergence}} + \sqrt{\frac{\text{Complexity}}{N}} + \underbrace{\Omega}_{\text{Constant}}, \text{(Eq. 2/3)}$$

[1] Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. 2010. A theory of learning from different domains. *Machine Learning*, 79(1-2), 151-175.

Brain Decoding and Inverse Inference in fMRI

Decoding the brain functions by using statistical learning methods [A brief introduction by Inria]



Domain Adaptation for Brain Decoding

Challenges for Machine Learning in Brain Decoding

- Small sample size
- High dimensional features
- Noisy

Our solution: public shared fMRI data + domain adaptation

- Neuroimaging data sharing (OpenfMRI/OpenNeuro, HCP...)
- Domain adaptation techniques

Data

OpenfMRI [1]: an open neuroimaging data sharing project [Dataset link].

Exp ID	Acc No.	Experiment Description	# Subjects
A	ds007	Stop signal with spoken pseudo word naming [2]	20
B	ds007	Stop signal with spoken letter naming [2]	20
C	ds007	Stop signal with manual response [2]	20
D	ds008	Conditional stop signal [3]	13
E	ds101	Simon task [Unpublished]	21
F	ds102	Flanker task [4]	26

Binary brain condition classification problem:

- "*Successful stop*" vs "*Unsuccessful stop*": A, B, C, and D
- "*Congruent correct*" vs "*Incongruent correct*": E and F
- Feature: brain voxels (feat level1 z-score maps)

[1] Poldrack, R. A.; Barch, D. M.; Mitchell, J.; Wager, T.; Wagner, A. D.; Devlin, J. T.; Cumba, C.; Koyejo, O.; and Milham, M. 2013. Toward open sharing of task-based fMRI data: the OpenfMRI project. *Frontiers in Neuroinformatics* 7:12.
[2] Xue, G.; Aron, A. R.; and Poldrack, R. A. 2008. Common neural substrates for inhibition of spoken and manual responses. *Cerebral Cortex* 18(8):1923–1932.
[3] Aron, A. R.; Behrens, T. E.; Smith, S.; Frank, M. J.; and Poldrack, R. A. 2007. Triangulating a cognitive control net-work using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *Journal of Neuroscience* 27(14):3743–3752.
[4] Kelly, A. C.; Uddin, L. Q.; Biswal, B. B.; Castellanos, F. X.; and Milham, M. P. 2008. Competition between functional brain networks mediates behavioural variability. *NeuroImage* 39(1):527–537.

Methodology

Assumption:
Predictions/decisions should be independent of experiments and subjects.

Theorem:

$$\epsilon_t(f) \leq \epsilon_s(f) + \underbrace{\hat{\rho}(f(\mathbf{X}), \mathbf{D})}_{\text{Statistical Dependence}} + \sqrt{\frac{\text{Complexity}}{N}} + \underbrace{\Omega}_{\text{Constant}}, \text{(Eq. 3/3)}$$

Practically, one-hot encoding for domain covariates $\mathbf{D}$

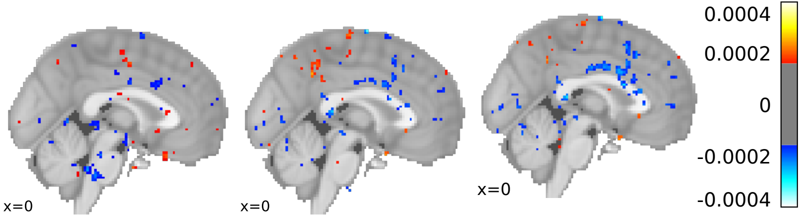
	Exp 1	...	Exp p	Sub 1	...	Sub q
Sample 1	1	0	0	1	0	0
Sample 2	1	0	0	1	0	0
Sample 3	0	1	0	0	1	0
...	...	...	...	...	...	...
Sample n	0	0	1	0	0	1

Results

	SVM	[1]	[2]	[3]	**Ours**
A→B	63.4±2.6	55.0±4.3	64.5±5.0	<u>65.7±2.2</u>	**78.6±2.9**
B→A	59.2±4.0	58.9±3.2	<u>71.2±3.5</u>	64.0±2.8	**79.6±2.7**
A→C	68.4±2.6	70.6±6.3	<u>78.4±2.9</u>	70.9±2.7	**87.4±2.1**
C→A	59.2±4.0	63.6±4.0	**70.6±4.0**	59.5±3.4	<u>68.8±3.8</u>
B→C	68.4±2.6	<u>86.4±5.3</u>	80.1±4.4	73.5±2.1	**90.5±1.5**
C→B	63.4±2.6	<u>75.3±4.0</u>	73.2±3.3	62.5±2.1	**77.4±2.8**
C→D	74.6±3.2	74.4±5.3	74.2±6.6	<u>81.3±4.0</u>	**87.1±2.3**
D→C	68.4±2.6	71.9±4.6	74.4±4.2	70.9±2.6	**87.4±2.6**
A→D	74.6±3.2	53.7±3.2	63.3±6.0	78.5±1.9	**86.2±1.9**
D→A	59.2±4.0	58.0±4.8	50.6±2.4	59.2±3.4	**67.1±3.8**
B→D	74.6±3.2	67.9±2.6	66.9±4.2	<u>85.6±2.3</u>	**94.2±1.9**
D→B	63.4±2.6	50.1±5.4	64.7±3.6	61.4±2.4	**73.7±3.1**
E→F	66.9±1.7	52.3±3.2	66.4±3.0	62.5±2.0	**74.0±4.0**
F→E	<u>53.3±2.8</u>	49.3±1.9	**53.7±4.5**	49.9±2.2	51.0±4.4
Avg	65.5±3.0	63.4±4.1	<u>68.0±4.2</u>	67.5±2.6	**78.8±2.8**
B&C→A	59.2±4.0	52.1±1.1	51.4±1.8	<u>63.2±2.5</u>	**80.3±3.1**
A&C→B	63.4±2.6	54.5±3.1	61.7±2.2	<u>67.6±1.6</u>	**79.5±2.0**
A&B→C	68.4±2.6	52.1±1.1	65.4±2.2	<u>71.9±2.6</u>	**89.9±1.7**
Avg	63.7±3.1	52.9±1.8	59.5±2.0	<u>67.6±2.2</u>	**83.2±2.3**

[1] Pan, S. J.; Tsang, I. W.; Kwok, J. T.; and Yang, Q. 2011. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks* 22(2):199–210.
[2] Yan, K.; Kou, L.; and Zhang, D. 2018. Learning domain invariant subspace using domain features and independence maximization. *IEEE Transactions on Cybernetics* 48(1):288–299.
[3] Long, M.; Wang, J.; Ding, G.; Pan, S. J.; and Philip, S. Y. 2013a. Adaptation regularization: A general framework for transfer learning. *TKDE* 26(5):1076–1089.

Model Coefficients Visualisation



(a) SVM (b) [1] (c) Ours

Figure: The top 1% learnt voxel weights in magnitude on A&B→C.

Consistent with the literature on response monitoring [2].

[1] Long, M.; Wang, J.; Ding, G.; Pan, S. J.; and Philip, S. Y. 2013a. Adaptation regularization: A general framework for transfer learning. *TKDE* 26(5):1076–1089.
[2] Braver, T. S.; Barch, D. M.; Gray, J. R.; Molfese, D. L.; and Snyder, A. 2001. Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral cortex*, 11(9): 825-836.

Summary

- Approach of leveraging public neuroimaging data
- Domain adaptation theorem and algorithm
- Experimental studies

Presented at the 34th AAAI Conference [Paper][Code].

One group (from Uni of Alberta and NUAA) followed our experimental setting and published a NeurIPS paper.

Thank you for your time

All questions are welcome

Reference:
[1] Zhou, S., Li, W., Cox, C., and Lu, H. 2020. Side information dependence as a regularizer for analyzing human brain conditions across cognitive experiments. In *AAAI* 34(04):6957-6964.
[2] Yousefnezhad, M., Selvitella, A., Zhang, D., Greenshaw, A., and Greiner, R. 2020. Shared Space Transfer Learning for analyzing multi-site fMRI data. In *NeurIPS*.