Organization for Human Brain Mapping

5841 Cedar Lake Road, Suite 204
Minneapolis, MN 55416 USA
Email: info@humanbrainmapping.org
Phone: (952) 646-2029
Fax: (952) 545-6073

View Abstract

Accounting for Voxel Neighbourhood Relationship in the SVM

D R Hardoon[1], J Mourao-Miranda[2], V Rocha Rego[3], J Shawe-Taylor[1]

[1]University College London, London, United Kingdom/[2]Kings College London, London, United Kingdom/[3]Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Introduction: In Neuroimage data analysis there are several preprocessing stages that must be applied before any statistical analysis can be done. The sMRI scan preprocessing procedures usually include transformation into standard space, segmentation and smoothing in space using a Gaussian filter. Despite the effectiveness of these preprocessing procedures, any incorrect step may seriously hamper the statistical analysis. Therefore, we propose to incorporate one of the preprocessing steps, namely the smoothing operator, into our neuroimage learning Support Vector Machine (SVM) through the use of a Gaussian prior on the weights. In effect we aim to learn the appropriate smoothing filter while learning the parameterisation for the neuroimage discrimination. We demonstrate our proposed novel technique on smoothed and non-smoothed sMRI in classifying patients and controls.

Methods:

The standard primal SVM includes a 2-norm constraint on the weight vector **w**. If we consider classifying brain scans with a linear kernel this encodes the expectation that the total activity across the whole brain is constrained. What it fails to capture is the expectation that activity levels form a smooth function across the brain. This is indirectly implemented by the smoothing in the preprocessing stage. An alternative approach is to put an appropriate prior over the weight vector that makes the machine learning biased towards weight vectors that are smooth. A natural way to define a prior is through a Gaussian Process (GP), that defines a multivariate Gaussian distribution P(**w**) over weight vectors **w**. Inclusion of log(P(**w**)) in the SVM objective will then bias the learner to pick more probable weights that are smoother. This will put larger values for pairs of variables that we expect to have similar values.

Results:

The aim of our experiments is to correctly classify patients versus controls in a Autistic Spectrum Disorder (ASD) and Post Traumatic Stress Disorder (PTSD) datasets. We use the SVMs as our learning algorithm together with the linear kernel on data that had been smoothed during preprocessing as well as data that had not been smoothed. Compared to our proposed SVM with Gaussian Process (SVM-GP) which is run on the non smoothed data. The PTSD results are given in Figure 1, whereas the results for ASD are given in Figure 2.

Conclusions:

Different disorders can effect the brain in different ways. Using a fixed smoothing preprocessing procedure for all diagnoses might not be ideal. In this work we propose a novel methodology of incorporating one of the preprocessing stages, the smoothing operation, into the learning function. We have demonstrated the feasibility of our proposed approach, SVM-GP, on two different sMRI datasets. Both have shown that we are able to significantly improve on the classification of SVM on non smoothed data, while achieving similar and better accuracy when compared to SVM on smoothed data. Furthermore we notice that as it is still possible to find, in the linear kernel scenario, the respective discriminating weight values.

Method	True-Positive	True-Negative	Accuracy
SVM (smoothed)	0.500	0.690	0.595
SVM (non smoothed)	0.630	0.440	0.535
SVM-GP covariance	0.630	0.440	0.535
SVM-GP ($\sigma = 2$)	0.630	0.440	0.535
SVM-GP ($\sigma = 4$)	0.560	0.440	0.500
SVM-GP ($\sigma = 6$)	0.690	0.630	0.660
SVM-GP ($\sigma = 8$)	0.630	0.630	0.630
SVM-GP ($\sigma = 10$)	0.630	0.630	0.630
SVM-GP Cross-validation on σ	0.560	0.630	0.595
Multi-kernel SVM-GP ($\mu = 0.5$)	0.690	0.630	0.660

Table 1. Results on post-traumatic stress disorder sMRI data, totalling 16 controls and 16 patients in a leave-subject-out routine. The results are given as an average. The proposed SVM-GP method is run on the non smoothed data. Futhermore we give the true-positive and true-negative accuracy rates are for comparison. We are able to observe that a number of SVM-GP configurations can outperform the results of SVM on smoothed data.

Method	True-Positive	True-Negative	Accuracy
SVM (smoothed)	0.700	0.740	0.720
SVM (non smoothed)	0.630	0.560	0.595
SVM-GP covariance	0.740	0.560	0.650
SVM-GP ($\sigma = 2$)	0.670	0.700	0.685
SVM-GP ($\sigma = 4$)	0.670	0.740	0.705
SVM-GP ($\sigma = 6$)	0.630	0.700	0.665
SVM-GP ($\sigma = 8$)	0.520	0.670	0.595
SVM-GP ($\sigma = 10$)	0.520	0.630	0.575
SVM-GP Cross-validation on σ	0.670	0.700	0.685
Multi-kernel SVM-GP ($\mu = 0.5$)	0.670	0.740	0.705

Table 2. Results on the Autistic spectrum disorder sMRI data, totalling 27 controls and 27 patients in a leave-subject-out routine. The results are given as an average. The proposed SVM-GP method is run on the non smoothed data. Furthermore we give the true-positive and true-negative accuracy rates are for comparison. We are able to observe that a number of SVM-GP configurations can outperform the results of SVM on non-smoothed data and is only 1.5% worse than the SVM on smoothed data.

References:

Argyriou, A (2005), 'Learning convex combinations of continuously parameterized basic kernels', *In Proceedings of COLT*, vol. , no. , pp. .

Ashburner, J (2000), 'Voxel-based morphometry - the methods', *NeuroImage*, vol. 26, no. 3, pp. 839-851.

Cristianini, N (2000), 'An Introduction to Support Vector Machines and other kernel-based learning methods', *Cambridge University Press*, vol. , no. , pp. .

Mourao-Miranda, J (2002), 'Classifying brain states and determining the discriminating activation patterns: Support Vector machine on function mri data', *NeuroImage*, vol. 28, no. 4, pp. 980-995.

Rasmussen, C (2006), 'Gaussian Processes for Machine Learning', *MIT Press*, vol. , no. , pp. .

Category: Modeling and Analysis
Sub-Category: Classification and predictive modeling