Visualization of nonlinear kernel models in neuroimaging by sensitivity maps

Rasmussen, Peter Mondrup; Hansen, Lars Kai; Madsen, Kristoffer Hougaard

Publication date:
2010

Visualization of nonlinear kernel models in neuroimaging by sensitivity maps

P. M. Rasmussen†*, K. H. Madsen‡, T. E. Lund*, L. K. Hansen†

†DTU Informatics, Technincal University of Denmark. ‡Danish Research Center for Magnetic Resonance, Hvidovre, Denmark. Center of Functionally Integrative Neuroscience, Aarhus, Denmark.

Aim

Recently there has been renewed interest in decoding information about mental states from distributed representational patterns in neuroimaging. It is highly desirable to identify in which brain locations the discriminative information resides. Here we investigate a model visualization for *nonlinear* kernel methods based on sensitivity analysis.

Materials and Methods

Data sets

Simulation: XOR coupling.

fMRI: Finger tapping experiment. 8 subjects. Block design.Two conditions: left and right hand finger tapping.10 rep. of each condition per subject.

Classification methods

- Support vector machine (SVM).
- Logistic regression (LogReg).
- Relevance vector machine (RVM).

Model evaluation

NPAIRS framework. Prediction accuracy and pattern reproducibility metrics.

Conclusion

Nonlinear classification models provide more flexibility than linear models. In addition to build a model with a good classification accuracy a competing goal is to interpret the model in terms of the neural representation that drives the predictions. The sensitivity map is a straightforward, yet powerful procedure for generation of meaningful global summary maps of nonlinear kernel methods.

References

[1] U. Kjems, L. K. Hansen, J. Anderson, S. Frutiger, S. Muley, J. Sidtis, D. Rottenberg, and S. C. Strother. The quantitative evaluation of functional neuroimaging experiments: Mutual information learning curves. *Neuroimage*, 15(4):772–786, 2002.

[2] Stephen LaConte, Stephen Strother, Vladimir Cherkassky, Jon Anderson, and Xiaoping Hu. Support vector machines for temporal classification of block design fmri data. *Neuroimage*, 26(2):317–329, Jun 2005.

[3] Michael E. Tipping. Sparse bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, June 2001.

Mathematical Modeling

Consider a labeled data set $\mathcal{D} = \{\mathbf{x}_n, t_n\}_{n=1}^{N}$, where $\mathbf{x}$ is a V dimensional feature vector that, in the context of classification in fMRI, contains a brain scan volume or some derived features, while t is the corresponding target variable. In a probabilistic framework the aim in supervised learning is to estimate the predictive distribution over t for a new input $\mathbf{x}$

$$p(t|\mathbf{x}, \mathcal{D}) = \int p(t|\mathbf{x}, \mathbf{w}) p(\mathbf{w}|\mathbf{x}, \mathcal{D}) d\mathbf{w}, \qquad (1)$$

where $\mathbf{w}$ are model parameters, while $p(t|\mathbf{x}, \mathbf{w})$ is the likelihood and $p(\mathbf{w}|\mathbf{x}, \mathcal{D})$ is the posterior distribution over the model parameters.

Binary classification: $p\left(t|\mathbf{x}, \mathbf{w}\right) = \sigma\left(t\mathbf{w}'\boldsymbol{\varphi}(\mathbf{x})\right).$

Linear model $y\left(\mathbf{x}\right) = \mathbf{w}'\boldsymbol{\varphi}(\mathbf{x})$, where basis functions φ may be *kernels*.

$\mathbf{k_x}$ holds the elements $k(\mathbf{x}_n, \mathbf{x})$, where n is the index of a specific training example. Examples are the linear kernel $k(\mathbf{x}_n, \mathbf{x}) = \mathbf{x}_n'\mathbf{x}$,
and the Gaussian kernel
$k(\mathbf{x}_n, \mathbf{x}) = \exp\left(-||\mathbf{x}_n - \mathbf{x}||^2/2\sigma^2\right)$.

Nonlinear classification

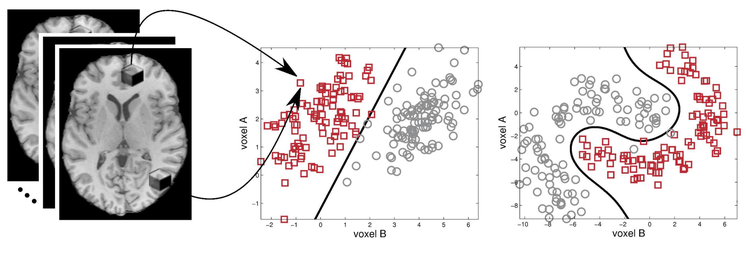

Figure 1: Analysis of spatial brain patterns by using a multivariate approach. Each scan is represented in terms of a high dimensional pattern vector (here only shown two dimensions denoted by voxel A and voxel B). In a classification scheme the goal is to assign the correct label/brain state to a new brain scan based on the spatial pattern vector. The symbols corresponds to two hypothetical brain states. In the middle panel the brain states are separable by a linear decision boundary, while a nonlinear decision boundary is required in the example in the left panel.

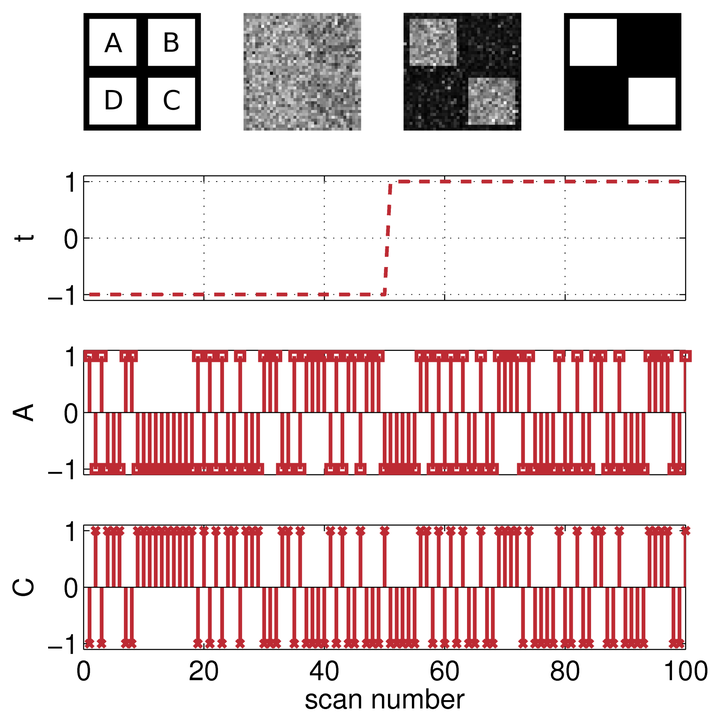

Figure 2: Simulation of regions that interact through a nonlinear relationship. The interrelation between the two regions is important to the classification task. This configuration is aimed at modeling a set of brain regions in which the interaction is modulated by the brain state. There are four spatially contiguous regions that can activate denoted by (A,B,C,D). Initially we let regions (A,B,D) be activated by random sequence taking values ± 1. The target signal, coding for an initial baseline taking values $t_n = -1$, and for an active state $t = 1$. The region (C) is activated with an XOR-sequence relative to region (A) and t_n, so that $C_n = A_n\, t_n$. The two regions (B,D) are unrelated to the stimulus reference function t_n but are included to ensure that the brain state can not be decoded based on regional measures, such as variance or non-Gaussianity. Additionally, Gaussian noise was added to the images to obtain a signal to noise ratio of 0.2. Classification accuracy with an SVM was 90% while the corresponding model with a linear kernel achieved 50% accuracy (corresponding to random guessing). Top panel left to right: simulation setup, example of input image, sensitivity map, "true map".

Model visualization

Usual visualization of *linear* kernel methods in functional neuroimaging

$$y\left(\mathbf{x}\right) = \mathbf{w}'\mathbf{k_x} = \mathbf{w}'\mathbf{X}\mathbf{x} \qquad (2)$$

where training examples are ordered in the rows of $\mathbf{X}$. $\mathbf{w}'\mathbf{X}$ is then considered as a discriminative map. We propose the sensitivity map as a general visualization of kernel methods in neuroimaging.

$$s_j = \int \left(\frac{\partial \log p(t|\mathbf{x}, \mathcal{D})}{\partial x_j}\right)^2 p(\mathbf{t}, \mathbf{x}) d\mathbf{t} d\mathbf{x}, \qquad (3)$$

where the visualization s_j resides in the same domain as the input data. In some models it is convenient to perform the integrals, in most cases we will approximate by empirical means over the training set $\mathcal{D}$. The sensitivity analysis or sensitivity map is a simple measurement of to what extent the predictive performance of a model depends on a given input.

For logistic regression

$$s_j = \int \gamma(\mathbf{w}, \mathbf{x}) \left(\frac{\partial \mathbf{w}'\mathbf{k_x}}{\partial x_j}\right)^2 p(\mathbf{x}) d\mathbf{x} \qquad (4)$$

where $\gamma(\mathbf{w}, \mathbf{x}) = \left(1 - \sigma\left(\mathbf{w}'\mathbf{k_x}\right)\right) \sigma\left(\mathbf{w}'\mathbf{k_x}\right)$.

Results - fMRI I

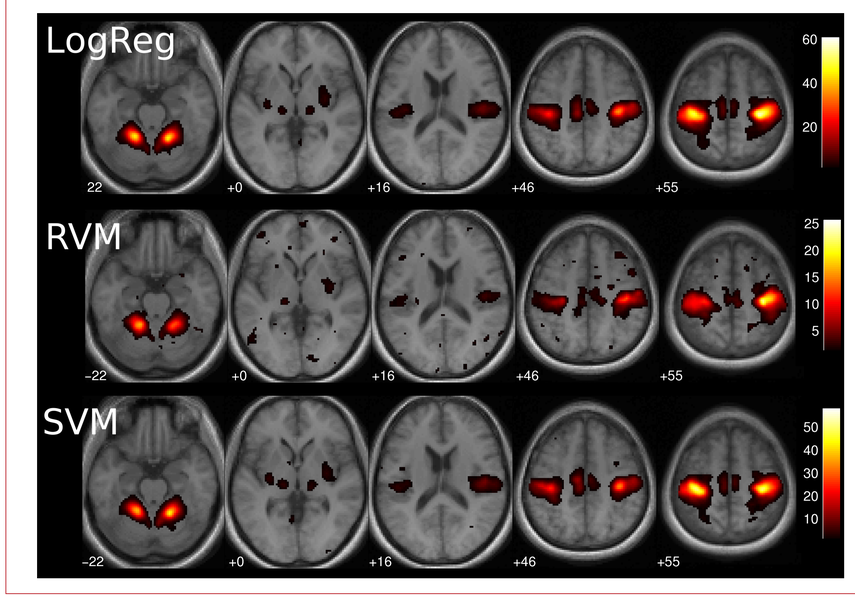

Figure 3: Classification of (Left) vs. (RIGHT) finger tapping with LogReg, RVM, and the SVM. Scans were represented in terms of the Gaussian kernel. The kernel width was set equal to the average input-space distance to the nearest 25 % points across all training examples. Model visualization was performed using the sensitivity map. The SVM showed highest prediction accuracy (99.1%), an average reproducibility of 0.86, and on average 64 support vectors out of 80 possible retained in the model. The LogReg model had a prediction accuracy (98.9%) and a reproducibility of 0.89. The RVM had a prediction accuracy of 98.0% and a reproducibility of 0.44. Furthermore, the RVM showed an exceptional degree of sparsity compared to the SVM with only three relevance vectors on average retained in the model. Models are visualized via reproducible sensitivity maps. The maps are arbitrary thresholded to show voxel in the upper 10 percentile.