

PRoNTo v3.1: Elastic Net Multiple Kernel Learning for Multisource Prediction

Janaina MOURÃO-MIRANDA^{1,2}, Konstantinos TSIRLIS³, Christophe PHILLIPS⁴

¹Centre for Artificial Intelligence, Department of Computer Science, University College London, UK.

²UCL Hawkes Institute, University College London, UK

³Department of Computer Science, University College London, UK

⁴GIGA-CRC Human Imaging, University of Liège, Liège, Belgium



Introduction

- The **Pattern Recognition for Neuroimaging Toolbox (PRoNTo)**¹ is an open-source MATLAB toolbox that brings machine learning to neuroimaging, with built-in support for classification and regression.
- A key feature of PRoNTo is its use of **regularised linear kernel methods**, which enable *computationally efficient, generalisable, and interpretable models* for high-dimensional data, even when the number of features exceeds the number of samples, a common situation in neuroimaging.
- Kernel methods represent data implicitly via a kernel function that encodes pairwise similarity; they are computationally efficient in high-feature low-sample-size settings because they operate on the kernel matrix ($\#samples \times \#samples$) rather than the raw data matrix ($\#samples \times \#features$).

Multisource integration:

- PRoNTo also integrates **multiple kernel learning (MKL)** to combine information from several kernels. Each kernel can represent a *different imaging modality* or a *different feature grouping* (e.g., a combination of regions of interest from a brain atlas).
- MKL learns an optimal weighted linear combination of these base kernels** under a chosen regularisation to maximize predictive performance. Common regularisation types include L1, which promotes sparsity; L2, which encourages smoothness; and elastic-net, which combines both to balance sparsity and grouping effects.
- We recently introduced a computationally efficient **Elastic-net MKL (ENMKL)** formulation that yields a simple analytical update for the kernel weights², see Figure 1.
- Here, we evaluated three approaches available in **PRoNTo v3.1** for combining multiple data sources for prediction: ENMKL², SimpleMKL (L1-MKL)³, and baseline models (SVM and KRR) trained on the unweighted sum of kernels.

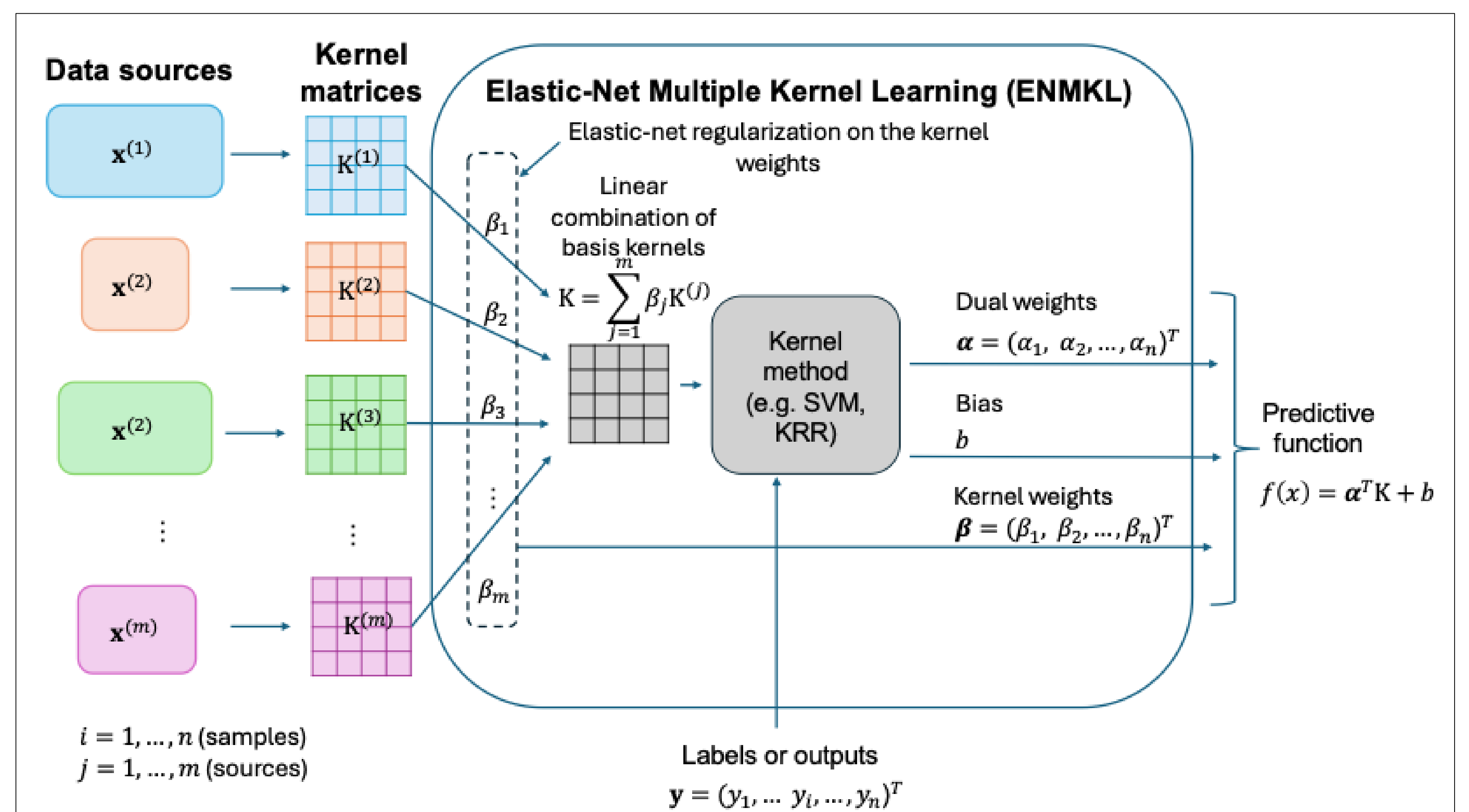


Figure 1: Elastic-net MKL algorithm

Methods

Datasets:

- Haxby dataset:**
 - Block-design fMRI visual paradigm in which participants passively viewed grayscale images from 8 categories: faces, cats, houses, chairs, scissors, shoes, bottles, and scrambled control images.
 - We used data from a single subject (participant 1): 12 runs, each with eight 24-second blocks.
 - The task was to classify pre-processed fMRI volumes recorded during face versus house viewing.
- IXI dataset:**
 - Structural MRI volumes from Guy's Hospital (102 subjects, ages 60–90).
 - The task was to predict each subject's age from their pre-processed structural MRI scans (GM density warped into MNI space).

Experimental Protocol

- For each dataset we used the Automated Anatomical Labeling (AAL)⁵ template to partition images into 116 anatomically defined regions.
- For each region we computed a linear kernel from the regional pattern of tissue or activation by using all voxels within that region.
- We compared (see Results and Figure 2):
 - two elastic-net MKL algorithm, Elastic-net Support Vector Machine (ENMKL-SVM) and Elastic-net Kernel Ridge Regression (ENMKL-KRR)
 - against L1-MKL classification and regression (SimpleMKL), and against baseline models trained on the unweighted sum of kernels, an SVM for classification and a KRR for regression.

Results

Haxby Results	Balanced Accuracy	AUC	Number of Kernels
ENMKL-SVM	97.22	100	56
L1-MKL class.	97.22	100	18
SVM	81.02	98	116

IXI Results	Mean Squared Error	Correlation	Number of Kernels
ENMKL-KRR	25.25	0.54	116
L1-MKL reg.	29.63	0.44	59
KRR	29.18	0.48	116

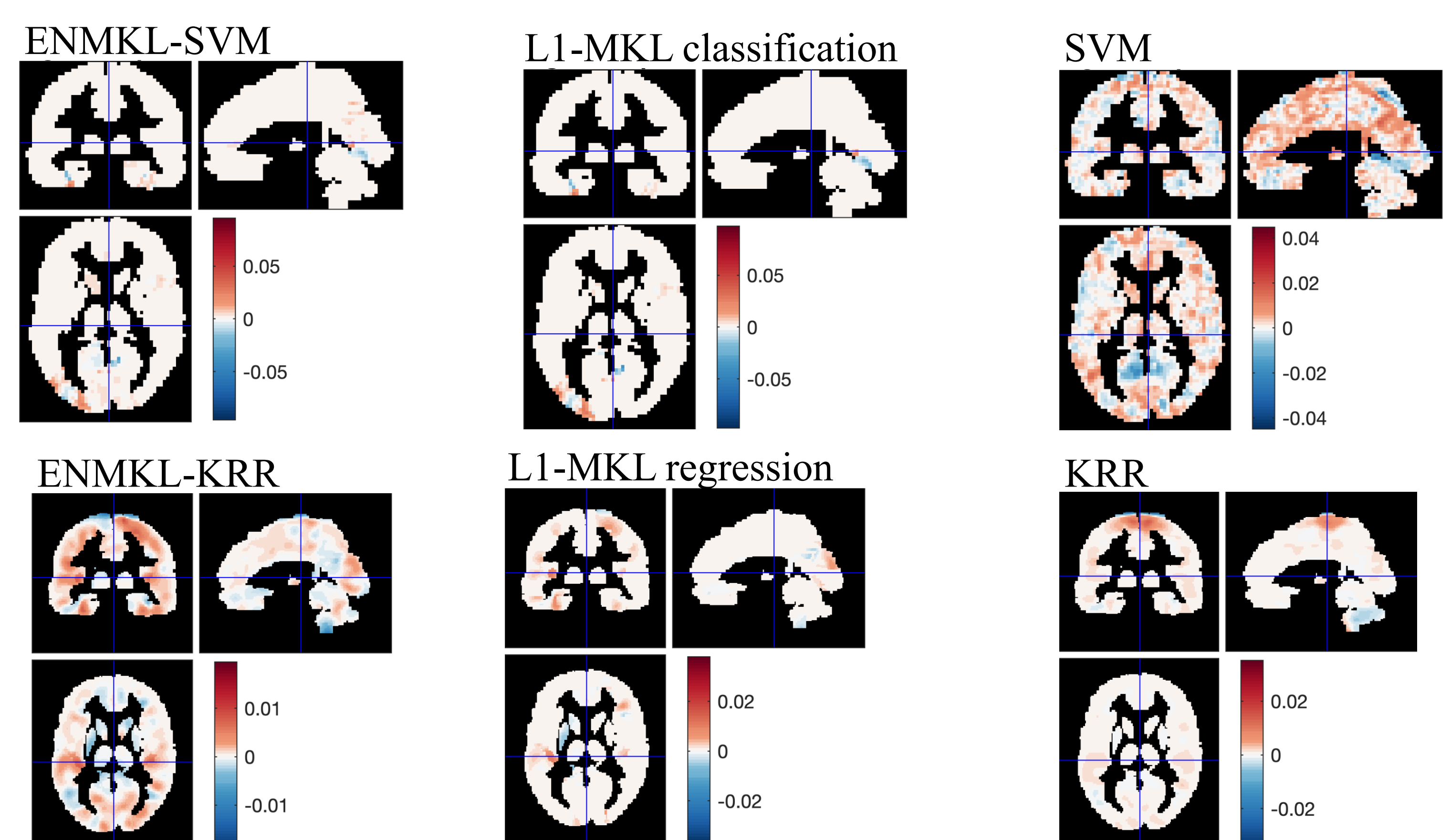


Figure 2: Voxel-wise weight maps for the different models and datasets

Conclusions

- Across both tasks, **ENMKL performed as well as or better** than L1-MKL and the baseline models.
- ENMKL can **identify sets of correlated regions** and reveal the effective level of sparsity in the predictive model, providing **enhanced interpretability**.

References

- Schrouff J et al. PRoNTo. *Neuroinformatics* 2013; 11:319–337.
- Mourão-Miranda J et al. Elastic-Net MKL. 2025; arXiv:2512.11547.
- Rakotomamonjy A et al. SimpleMKL. *J Mach Learn Res* 2008;9 :2491–2521.
- Haxby JV et al. Faces vs objects. *Science* 2001; 293:2425–2430.
- Tzourio-Mazoyer N et al. AAL atlas. *NeuroImage* 2002; 15:273–289.