Kinds of Ensembles

Tested on . . . dataset

David Nardi

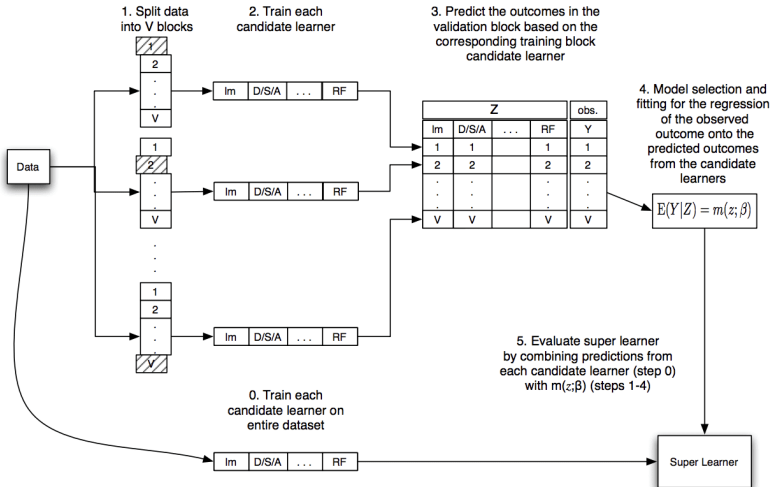June 11, 2024

MSc in AI, University of Florence

Binary classification task

Class distribution: 0:0.49 – 1:0.51

Approaches

- Decision tree

- Random forest

- AdaBoost

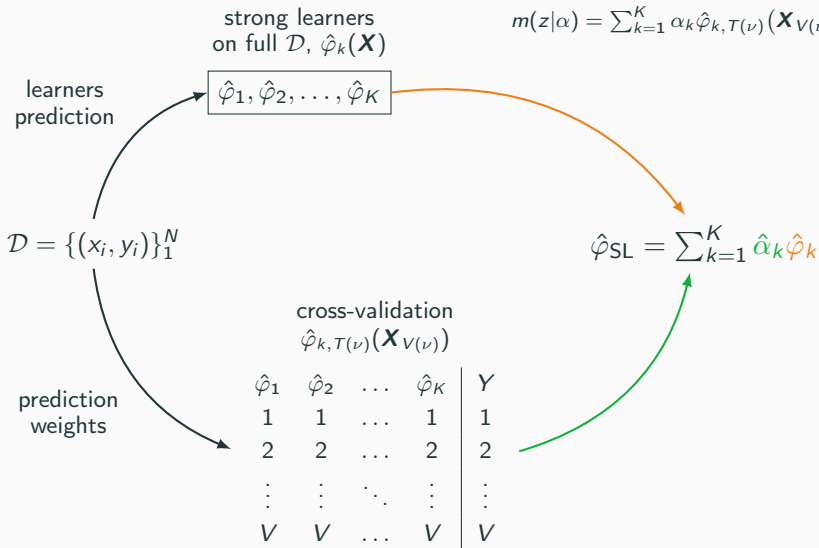- Ridge logistic regression

- Super learner

What's a Super Learner?

1. Split data into V blocks
2. Train each candidate learner
3. Predict the outcomes in the validation block based on the corresponding training block candidate learner
4. Model selection and fitting for the regression of the observed outcome onto the predicted outcomes from the candidate learners

$$E(Y|Z) = m(z; \beta)$$

5. Evaluate super learner by combining predictions from each candidate learner (step 0) with $m(z; \beta)$ (steps 1-4)

0. Train each candidate learner on entire dataset

The Super Learner flow diagram

$\hat{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{N} L(Y_i, m(z_i|\alpha))$

$m(z|\alpha) = \sum_{k=1}^{K} \alpha_k \hat{\varphi}_{k, T(\nu)}(\boldsymbol{X}_{V(\nu)})$

strong learners
on full $\mathcal{D}$, $\hat{\varphi}_k(\boldsymbol{X})$

$\boxed{\hat{\varphi}_1, \hat{\varphi}_2, \ldots, \hat{\varphi}_K}$

learners
prediction

$\mathcal{D} = \{(x_i, y_i)\}_1^N$

$\hat{\varphi}_{\text{SL}} = \sum_{k=1}^{K} \hat{\alpha}_k \hat{\varphi}_k$

cross-validation
$\hat{\varphi}_{k, T(\nu)}(\boldsymbol{X}_{V(\nu)})$

prediction
weights

$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\ldots$	$\hat{\varphi}_K$	Y
1	1	$\ldots$	1	1
2	2	$\ldots$	2	2
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
V	V	$\ldots$	V	V

Model	Train score	Test score
CART	0.0000	0.0000
Random forest	0.0000	0.0000
AdaBoost	0.0000	0.0000
Super learner	0.0000	0.0000

📄 E. C. Polley, and M. J. van der Laan
Super Learner in Prediction
U.C. Berkeley Division of Biostatistics Working Paper Series.
Working Paper 266, 2010