

Kinds of Ensembles

Tested on ...dataset

David Nardi

June 11, 2024

MSc in AI, University of Florence

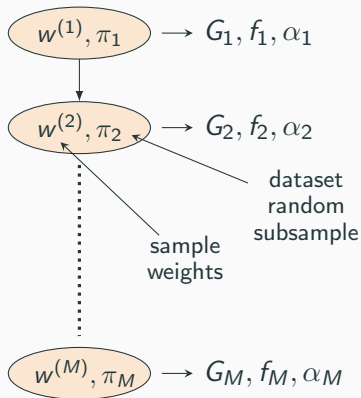
Binary classification task

Class distribution: 0:0.49 – 1:0.51

Approaches

- Decision tree
- Random forest
- AdaBoost
- Ridge logistic regression
- Super learner

AdaBoost algorithm



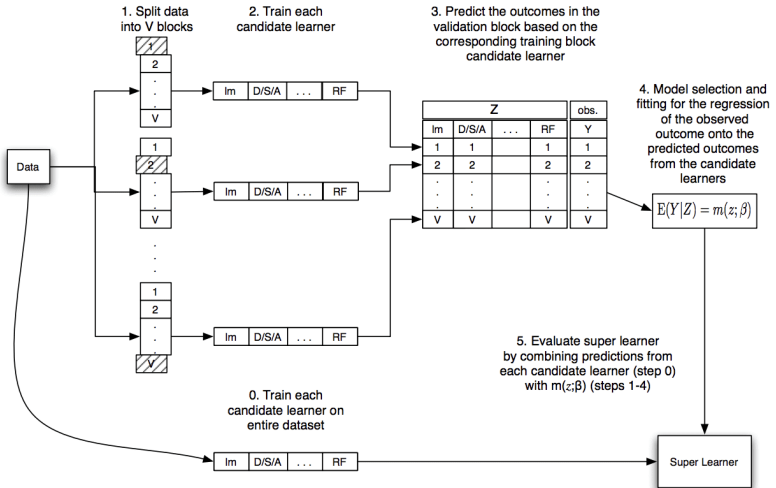
```
ada::ada(formula, data,
  loss="exponential",
  type="discrete",
  iter, ← M
  nu=0.1, ← λ
  bag.frac=0.5 ← π
)
```

$$f_m(x) = f_{m-1}(x) + \lambda \alpha_m G_m(x)$$

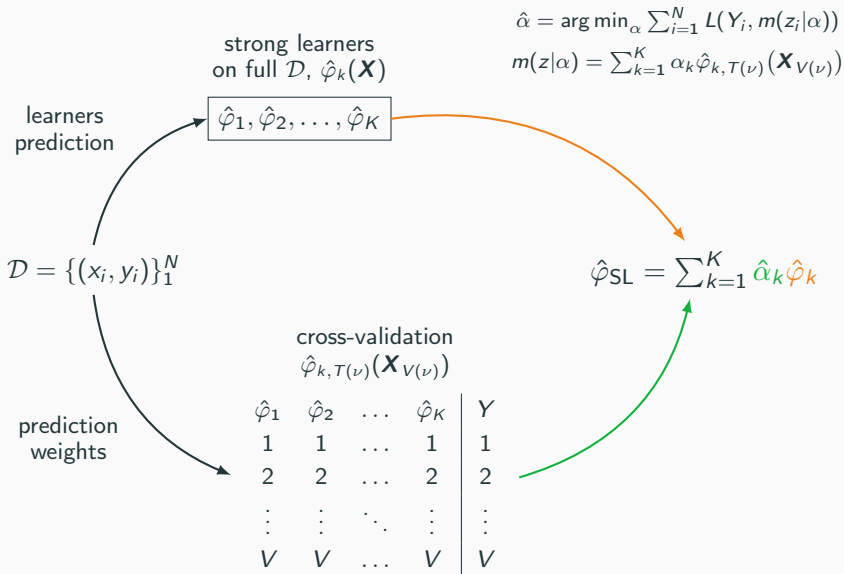
$$G(x) = \text{sign}(f_M(x))$$

$$L(y, f(x)) = \exp(-yf(x))$$

What's the Super Learner?



The Super Learner flow diagram



Package SuperLearner

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