Ensembling

Quiz, 6 questions

2 points

1.

Suppose we are given a train set and test set, that came from the same distribution. We want to use stacking and choose between the validation schemes described in the reading material.

Select the true statements about the validation schemes.

- [x] Scheme **e)** gives the validation score with the least variance, if compared to schemes **a)** -- **d)**.
- [x] Scheme **d)** is less efficient from computational perspective than scheme **a).** That is, if a dataset is very large, scheme **a)** is usually preferred over scheme **d)**.
- [ ] Scheme **d)** is more efficient from computational perspective than scheme **a).** That is, if a dataset is very large, scheme **d)** is usually preferred over scheme **a)**.

---

2 points

2.

Definition: we will call a validation scheme *fair* if the set, that we use to validate meta-models comes from the same distribution as the meta-test set. In other cases we will call validation scheme *leaky*. In other words in a *fair* validation scheme the set that we use to validate meta-models was not used in any way during training first level models.

Select *fair* validation schemes. The definition for the schemes can be found in the reading material.

- [x] **d)** Holdout scheme with OOF meta-features
- [x] **e)** KFold scheme with OOF meta-features
- [ ] **b)** Meta holdout scheme with OOF meta-features
- [x] **a)** Simple holdout scheme
- [ ] **c)** Meta KFold scheme with OOF meta-features

---

1 point

3.

Which of the following ensembling methods can potentially learn "conditional averaging" (video 1)?

- [ ] Weighted average
- [x] Boosting on trees
- [ ] Bagging
- [x] Stacking

1 point

4.

The benefits of the weighted average compared to more advanced ensembling techniques is that

- [x] It is faster to implement and to run
- [ ] It usually gives better quality
- [x] It is less prone to overfitting

1 point

5.

In general case, which set of base models is probably the best for stacking?

- ( ) [Random Forest, Extra Trees Classifier, GBDT, RGF]
- ( ) [Logistic Regression, SVM, Random Forest, Extra Trees Classifier, GBDT]
- (•) [SVM, GBDT, Neural Network, kNN]
- ( ) [kNN, SVM, Logistic Regression, Neural Net]

2 points

6.

Ensembling

Quiz, 6 questions

Suppose we are given a classification task. In a simple two model linear mix we usually use weights α for the first model and β for the second one. The coefficients are usually chosen such that $\alpha + \beta = 1$, because convex combination of probability vectors is a probability vector.

Still, sometimes it is beneficial to tune α and $beta$ independently, e.g. mix with $\alpha = 0.1$ and $\beta = 0.8$ works best.

However, for some metrics it never makes sense to tune α and β independently. That is, searching for independent α and β will never give you better results than searching for weights, constrained to be $\beta = 1 - \alpha$. Select such metrics.

- [ ] LogLoss
- [ ] Hinge loss
- [x] Accuracy (implemented with argmax)
- [x] AUC

I, **Jiadai Zhao**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Learn more about Coursera's Honor Code

Submit Quiz