Microsoft: DAT203x Data Science and Machine Learning Essentials

QUESTION 5 (1/1 point)

Which two of the following approaches can you use to determine which features to prune in an Azure ML experiment?

- [x] Use the Permutation Feature Importance model to identify features of near-zero importance. ✔
- [ ] Use the Cross Validation module to identify folds which indicate the model does not generalize well.
- [x] Prune features one at a time to find features which reduce model performance or have no impact on model performance as measured with the Evaluate Model module. ✔
- [ ] Use the Split module to create training, test and evaluation data sub-sets to evaluate model performance.

✔

Note: Make sure you select all of the correct options—there may be more than one!

EXPLANATION

The goal of pruning is to eliminate features from the dataset which either reduce model performance or have no impact on model performance but increase complexity and may reduce how well the model generalizes. Feature importance can be useful guide to finding feature pruning candidates. Once a feature has been pruned from a dataset, the performance of the model must be measured to assess the actual effect.

You have used 1 of 2 submissions

Module 4 Review
Homework due Oct 30, 2015 at 00:00 UTC

- Module 5: Recommenders and Publishing Your Work
- Final Exam