XGBoost
Part 5: Attribute importance in XGBoost

By: Noureddin Sadawi, PhD

University of London

Feature (i.e. attribute) importance

- A score that is used to indicate how useful each feature (or input variable) is.
- Think of coefficient values in regression.
- How much of a role does that feature play in deciding the outcome?
- This is done based on the data that is being used for model training.
- Features can be ranked based on their importance.

Feature importance in XGBoost

- XGBoost creates a number of decision trees internally.
- The relative importance of a feature increases as it is used more in making key decisions.
- In a single decision tree, the importance of a feature can be inferred from the performance improvement made at split point based on that feature.
- The average feature importance(s) can be computed from all decision trees created by XGBoost (e.g. the average importance of each feature can be calculated).

Feature importance in XGBoost

- 'Feature importance is defined only for tree boosters.
- Feature importance is only defined when the decision tree model is chosen as base learner (booster=gbtree).
- It is not defined for other base learner types, such as linear learners (booster=gblinear).'

  https://xgboost.readthedocs.io/en/latest/python/python_api.html

The get_score function in XGBoost

'Get feature importance of each feature. Importance type can be defined as:

- 'weight': the number of times a feature is used to split the data across all trees.
- 'gain': the average gain across all splits the feature is used in.
- 'cover': the average coverage across all splits the feature is used in.
- 'total_gain': the total gain across all splits the feature is used in.
- 'total_cover': the total coverage across all splits the feature is used in.'

https://xgboost.readthedocs.io/en/latest/python/python_api.html

Example feature importance plots