4.05 Hyperparameter Tuning

Hyperparameters

- Built-in quantities of models that we can use to fine-tune our results. For example what value of k do we select?

- These are quantities our model cannot learn... we must decide on these ourselves!

- These are different from statistical parameters, which are quantities a model can learn.

- Different values for hyperparameters can result in substantially different models.

GridSearchCV

- Allows you to identify ==“optimal values”== by defining a grid of hyperparameters (params) that will be searched using k-Fold Cross Validation

- Runs through all the param values that are fed into the param grid and produces best combination of params, based on scoring metric
  - Param 1 (n_neighbors | n = 10 to 50)
  - Param 2 (weights | weights = uniform or distance)
  - ….
  - Param Z (……………)

GridSearchCV

- What to input into GridSearch CV?

  - estimator: Model being used

  - param_grid: dictionary that contains all of the parameters to try

  - cv: cross-validation, the number of cv folds for each combination of parameters

  - scoring [optional]: evaluation metric to use when ranking results