Random Forest

For Decision Tree Regression

Definitions

- **Random Forest** is a version of ensemble learning, we have other versions of ensemble learning such as gradient boosting.
- **Ensemble learning** is when we take multiple algorithms or the same algorithm multiple times and put them together to make something much more powerful than the original.
  - The wisdom of the crowds

Random Forest Steps for Decision Trees Regression:

- Step 1: pick at random K data points from the training set.
- Step 2: build a decision tree based on these K data points.
- Step 3: choose the number of trees you want to build and repeat Steps 1 & 2
- Step 4: when trying to predict a new data point, make each one of the trees predict the value of Y for the data point in question, and assign the new data point the average across all of the predicted Y values.

Random Forest

For Decision Tree Classification

Random Forest Steps for Decision Trees Classification:

- Step 1: Pick at random K data points from the training set
- Step 2: build a Decision Tree based on these K data points
- Step 3: choose the number of trees you want to build and repeat steps 1 & 2
- Step 4: when trying to predict a new data point, make each one of your trees predict the category to which the data point belongs, and assign the new data point to the category that wins the majority vote