* Bagging is a special case of random forests under which case?

*When the subset we choose is equal to the size of the original dataset. When m equals p.*

* What are the hyperparameters we can control for random forests?

*We can control tree size, number of trees to take and number of features to split.*

* Suppose you have the following paired data of (x,y): (1,2), (1,5), (2,0). Which of the following are valid bootstrapped data sets? Why/why not?
  1. (1,0), (1,2), (1,5)
  2. (1,2), (2,0)
  3. (1,2), (1,2), (1,5)

*(1,2), (1,2), (1,5) is the valid bootstrap data set example because it contains subset of original data set and there are three subsets which is equal to the size of the original data set. Choice b does not have equal size as that of original data set. Choice (a) has a subset (1,0) which is not a proper subset of original data set.*

* For each of the above valid bootstapped data sets, which observations are out-of-bag (OOB)?

Subset ( 2,0) is the out of bag observation .

* You make a random forest consisting of four trees. You obtain a new observation of predictors, and would like to predict the response. What would your prediction be in the following cases?
  1. Regression: your trees make the following four predictions: 1,1,3,3.

Prediction will be 2 as it is the average of 1 and 3

* 1. Classification: your trees make the following four predictions: “A”, “A”, “B”, “C”.

The prediction will be A as A is the mode of this sample.