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

Random Forest is an improvement over the Bagging. Bagging can have structural similarities and tend to be correlated. Averaging highly correlated trees doesn’t reduce the large variance. To avoid this correlation, Random Forest algorithm changes the learnt sub trees accordingly so that the resulting predictions will have a lesser correlation.

When n= p.

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

Length of the decision tree, Number of Predictors

* **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)-** Invalid
  2. **(1,2), (2,0)-**Invalid
  3. **(1,2), (1,2), (1,5)-**Valid
* **For each of the above valid bootstapped data sets, which observations are out-of-bag (OOB)?**

(2,0)

* **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.**

2- Average of All

* 1. **Classification: your trees make the following four predictions: "A", "A", "B", "C".**
     1. More number of predictions