1. What is a decision tree, how do we decide how to make splits?

A decision tree is a hierarchical model for supervised learning whereby the local region is identified in a sequence of a recursive split in a smaller number of steps. In essence, we are looking for a hierarchical decision based on predictor data.

The decision to make a split is determined by quantifying the impurity measure.  The split is pure if after the split, for all branches, all the instances choosing a branch belonging to the same class.

1. How can we create a set of rules based on a decision tree?

The goal of a rule-based decision tree is to generate a decision tree from a set of rules instead of data. When building a decision tree from rules, the method assigns attributes to the nodes using criteria based on the properties of the attributes in the decision rules, rather than statistics regarding their coverage of the data examples.

1. What is the difference between pre-pruning and post-pruning?

Pruning means to compress, optimized, and simplify a decision three.  We remove subtrees for better generalization (decrease variance).  In Prepruning: Early stopping

Post pruning: Grow the whole tree then prune subtrees that overfit on the pruning set.

Prepruning is faster, post-pruning is more accurate (requires a separate pruning set)

1. What is the main reason to use ensembles of learners?

Ensemble methods use multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone.

1. What is bagging?

Bagging or Bootstrap Aggregation is used when our goal is to reduce the variance of a decision tree.  The idea is to create several subsets of data from the training sample chosen randomly with replacement.

1. What is boosting?

Boosting is another ensemble technique to create a collection of predictors. In this technique, learners are learned sequentially with early learners fitting simple models to the data and then analyzing data for errors.