

1. What are the advantages and disadvantages of decision trees compared to linear models presented in the Chapter 6?

Advantages:

- Trees are easy to explain because of their outputs. They are more understandable.
- It's close to method of human decision making.
- They can be displayed as graphs.
- They can handle qualitative predictors.

Disadvantages:

- Less accuracy than regression and classification approaches.
- Non-robust. Small change in data may causes the large change on output of decision tree.

(Page: 315)

2. How does Bagging solve the problem of high variance in decision trees?

Bagging basically taking repeated samples from the (single) training data set -bootstrap- and averaging a set of observations to reduce variance. In other words, bagging generates different bootstrapped training datasets. Then, train the method on these different datasets. Final part, average all predictions.

For high variance in decision trees, when bagging is applied, it constructs N regression trees using bootstrapped train datasets, and average the resulting predictions. These trees are deep and not pruned. Each tree has high variance but low bias. Averaging these trees reduce variance.

(Page 316)

3. How can a strong predictor cause high variance in Bagging and how do random forests solve this problem?

Strong predictor will cause to fail on Bagging because all generated trees try to split with this strong predictor and all trees will be correlated.

Random forest overcome the this problem by decorrelating the generated trees. RF forces each split to consider only a subset of the predictors.

(Page: 320)

4. How does Boosting work?

Boosting works similar with bagging but the trees are grown sequentially on boosting; each tree grown using information from previously grown trees. It learn from previous trees !

(Page 321)