

Questions for Review

Decision Trees

CSC357 Advanced Topic: Machine Learning

28 January 2020

1. Look at the first paragraph in the section titled *Instability* in *Chapter 6: Decision Trees*. What are some of the attractive qualities of Decision Tree Models?

Perform both classification and regression tasks, simple to understand and interpret, easy to use , versatile and powerful.

2. Scan the chapter. What are some of the hyperparameters that we might use in our construction of a Decision Tree model?

`max_depth`, `min_samples_split`, `min_samples_leaf`, `min_weight_fraction_leaf`, `max_leaf_nodes`, and `max_features`

3. What is the *root node* of a tree? What is a *leaf node*? What other kinds of nodes are there?

A tree has a single root node. The root node has no parent.
A leaf is a node with no children (no descendants).
An internal node has both a parent and at least one child.

4. What is the Graphviz command that will produce a .jpg file from a .dot file?
-

`dot -t jpg my-tree.dot -o my-tree.jpg`

5. What kind of preparation of data is necessary a building a decision tree model?
-

not too much

6. What is a *white box model*?
-

WRITE YOUR ANSWER HERE.

7. What is the relationship between the *gini* attribute and a node's *impurity*?

WRITE YOUR ANSWER HERE.

8. What is an alternative to the *gini* measure?

WRITE YOUR ANSWER HERE.

9. What is the *CART algorithm*?

WRITE YOUR ANSWER HERE.

10. What is a *greedy algorithm*?

WRITE YOUR ANSWER HERE.

11. What is an *NP-Complete algorithm*?

WRITE YOUR ANSWER HERE.

12. What is the computational complexity of making a prediction with a Decision Tree? What is the computational complexity of training a Decision Tree? (Be clear about what m and n denote.)

WRITE YOUR ANSWER HERE.

13. Distinguish between *parametric* and *nonparametric* models?

WRITE YOUR ANSWER HERE.

14. Géron shows how a Decision Tree can be used for regression. How does the tree for regression differ from the tree for classification?

WRITE YOUR ANSWER HERE.

15. What is the significance of *orthogonal decision boundaries* in Decision Tree models?

WRITE YOUR ANSWER HERE.

16. Instability is common in Decision Tree models. What does that mean?

WRITE YOUR ANSWER HERE.
