

1. These are all characteristics of decision trees, EXCEPT:

1 / 1 point

- ☐ They split nodes into leaves
- ☒ They have well rounded decision boundaries
- ☐ They segment data based on features to predict results
- ☐ They can be used for either classification or regression

✓ **Correct**

Correct! For more information please review the *Introduction to Decision Trees* lesson.

2. Decision trees used as classifiers compute the value assigned to a leaf by calculating the ratio: number of observations of one class divided by the number of observations in that leaf E.g. number of customers that are younger than 50 years old divided by the total number of customers.

1 / 1 point

How are leaf values calculated for regression decision trees?

- ☐ mode value of the predicted variable
- ☐ weighted average value of the predicted variable
- ☐ median value of the predicted variable
- ☒ average value of the predicted variable

✓ **Correct**

Correct! For more information please review the *Decision TreeSplit* lesson.

3. These are two main advantages of decision trees:

1 / 1 point

- ☐ They do not tend to overfit and are not sensitive to changes in data
- ☐ They are resistant to outliers and output scaled features
- ☒ They are very visual and easy to interpret
- ☐ They output both parameters and significance levels

✓ **Correct**

Correct! For more information please review the lesson Pros and Cons of Decision Trees.

4. How can you determine the split for each node of a decision tree?

1 / 1 point

- ☐ Find the split that induces the largest entropy.
- ☐ Randomly select the split.
- ☒ Find the split that minimizes the gini impurity.
- ☐ Use a nonlinear decision boundary to find the best split.

✓ **Correct**

Correct. The split for each node is determined by the split that minimizes the gini impurity or the entropy.

5. Which of the following describes a way to regularize a decision tree to address overfitting?

1 / 1 point

- ☒ Decrease the max depth.
- ☐ Increase the max depth.
- ☐ Reduce the information gain.
- ☐ Increase the number of branches.

✓ **Correct**

Correct. Decreasing the max depth can help to prune the tree. Any impure leaves can be assigned to the majority class.

6. What is a disadvantage of decision trees?

1 / 1 point

- ☐ Scaling is required.
- ☒ They tend to overfit.
- ☐ They can get too large.
- ☐ They are difficult to interpret.

✓ **Correct**

Correct. Decision trees tend to add high variance; that is, they tend to overfit.

7. What method can you use to minimize overfitting of a machine learning model?

1 / 1 point

- ☐ Choose the hyperparameters that maximize goodness of fit on your training data.
- ☒ Tune the hyperparameters of your model using cross-validation.
- ☐ Decrease the variance of your test data.
- ☐ Increase the variance of your training data.

✓ **Correct**

Correct. As you've done with other algorithms, you want to tune your hyperparameters using cross-validation. This means choosing the hyperparameters that maximize the estimated goodness of fit on unseen test data.

8. Concerning Classification algorithms, what is a characteristic of K-Nearest Neighbors?

1 / 1 point

- ☒ Training data is the model
- ☐ The model is just parameters
- ☐ Fitting can be slow
- ☐ Prediction is fast

✓ **Correct**

Correct. For K-Nearest Neighbors the training data is the model.

9. Concerning Classification algorithms, what are the characteristics of Logistic Regression?

1 / 1 point

- ☐ The training data is the model, fitting is fast, predicting class for new records can be slow, and the decision boundary is flexible
- ☒ The model is just parameters, fitting can be slow, prediction is fast, and the decision boundary is simple and less flexible
- ☐ The training data is the model, fitting is fast, prediction is fast, and the decision boundary is flexible
- ☐ The model is just parameters, fitting is fast, prediction is fast, and the decision boundary is flexible

✓ **Correct**

Correct.

10. When evaluating all possible splits of a decision tree what can be used to find the best split regardless of what happened in prior or future steps?

1 / 1 point

- ☒ Greedy Search
- ☐ Classification
- ☐ Regularization
- ☐ Logistic regression

✓ **Correct**

Correct. Greedy Search is used to find the best split regardless of what happened in prior or future steps.