



✓ **Congratulations! You passed!**  
TO PASS 80% or higher

Keep Learning

GRADE  
85.71%

## Understanding Classification with Decision Trees and k-NN

LATEST SUBMISSION GRADE  
85.71%

1. What method does scikit-learn use for classifying operational data?

0 / 1 point

fit(X,y)

! Incorrect

Review Lecture 1, Video 2: Classification in scikit-learn to remind yourself how it works.

2. What is the decision tree learning algorithm trying to do at each node in the tree?

1 / 1 point

- ☒ Create the split that makes the biggest difference in the resulting data set.
- ☐ Create the split that minimizes the difference in the resulting sets.
- ☐ Split the data to achieve complete separation in nodes.
- ☐ Find a binary question that tells you whether an email is spam.

✓ Correct

Correct! The learning algorithm is looking for splitting measures that creates the most separation between nodes.

3. What does it mean if your model has **overfit** the data?

1 / 1 point

- ☐ It hasn't captured enough detail from the training data about the question.
- ☐ It has captured details in the test data that are irrelevant to the question.
- ☐ It hasn't captured enough detail from the test data about the question.
- ☒ It has captured details in the training data that are irrelevant to the question.
- ☐ It has memorized the correct answers to the test.

✓ Correct

Correct! A model overfits the data when it is matching too closely to all the details in the training data.

4. If we are measuring the distance between three points (A, B, and C), and distance from A to B is 5 units and the distance from B to C is 6 units, what else might be true?

1 / 1 point

- ☒ AC = 10 units
- ☐ CB = 4 units
- ☐ CA = 12 units
- ☐ AB = -8 units
- ☐ Nothing because it depends on what distance function you're using.

✓ Correct

Correct! It is possible for the distance from A to C to be 10 units long.

5. What do you need to keep in mind when picking a "k" for k-Nearest Neighbours?

1 / 1 point

- ☐ The number should be large to prevent bias.
- ☐ The number should be four.
- ☒ The number needs to be chosen carefully when there are three or more classes.

✓ **Correct**

True! It is important to choose a k that minimizes ties, which is particularly difficult when considering more than binary classification.

- ☒ The number shouldn't be too small, to prevent influence from local, minute variation.

✓ **Correct**

True! A small k makes the model especially susceptible to noise.

- ☐ The number should be odd to prevent ties.
- ☒ The number shouldn't be too big, to prevent influence from very dissimilar points.

✓ **Correct**

True! A large k means considering very distant, and therefore dissimilar, points when determining the class of an example.

- ☐ The number doesn't matter that much and you can use whatever you feel like.

6. What makes classification different from regression? Select all that apply.

1 / 1 point

- ☒ Labels are categories

✓ **Correct**

Correct! Regression problems have numbers as labels.

- ☐ Labels must be supplied by a human supervisor.
- ☐ Regression builds a QuAM
- ☐ Classification does not require labels
- ☒ Labels form an unordered set

✓ **Correct**

Correct! Unlike regression problems, class labels do not have numeric meaning.

7. How do you know when your learning algorithm has overfit a model?

1 / 1 point

- ☐ The performance is good.
- ☒ The training error is low but the test error is high.
- ☐ The training error and test error are high.
- ☐ The test error is low
- ☐ The training error is low.

✓ **Correct**

Correct! Overfitting occurs when your model performs well on the training data but poorly on the test data.