



[Course](#) > [Evaluat...](#) > [Dive D...](#) > [Knowle...](#)

## Knowledge Checks

m7q1

1.0/1.0 point (graded)

After being hired by Robert Downey Sr's start-up "SpaceZ", you've been placed in charge of the artificial intelligence that controls the HyperExplorer Pod. This vessel will travel at very fast velocities through the asteroid belt to explore the long-term effects of low-gravity and space-travel on the humans, animals, and plants onboard. It's important the craft is able to maneuver around asteroids safely, but gas clouds aren't as important to avoid (we know gas clouds don't just hang out in space, but work with us here...) You have around 90-thousand data samples of various space instrument readings stored as features, all of which were collected from earlier satellite launches. The first item of business you need to take care of is identifying the type of objects the vessel encounters, such as comets, asteroids, aliens, gas clouds, and solar flares.

Which of the following algorithms makes the most sense to use to do this?

☐ Decision Tree Regression

☐ Linear Regression

☐ Randomized PCA

☐ K-Means

☒ SVC



Submit

You have used 1 of 2 attempts

## m7q2

1.0/1.0 point (graded)

Given the column definitions: Alligator, Crocodile, and Water Snake, and the following confusion matrix: Which target is your model doing a really bad job of correctly predicting, and what class does it actually think that target is?

3   0   1
4   1   1
0   2   3

☒ Crocodile, Alligator☐ Alligator, Crocodile☐ Crocodile, Water Snake☐ Water Snake, Alligator☐ Alligator, Water Snake

You have used 1 of 2 attempts

## m7q3

1.0/1.0 point (graded)

If all you cared about was how *fast* an algorithm could come up with a decent approximation of how much wool a sheep produces per year given the altitude above sea-level it lives, how much it eats per day, and the amount of day-light hours it receives each month, and the existence (or not) of a few genetic markers...

Which of the following algorithms are you most likely to use?

☐ Logistic Regression☒ Linear Regression

☐ Boosted Decision Trees Regression

☐ Isomap

☐ Neural Networks

Submit

You have used 1 of 2 attempts

m7q4

1.0/1.0 point (graded)

If a rocket ship's navigation computer needs to classify asteroids to avoid them...

☐ True positives are the most harmful, because an android will certainly strike the ship.

☐ True negatives are the most harmful, because the model failed to classify the asteroid, causing it to strike the ship.

☐ False positives are the most harmful, because the model got confused resulting in the ship getting struck.

☒ False negatives are the most harmful, because the model didn't tell the ship to dodge the asteroid.



Submit

You have used 1 of 2 attempts

m7q5

1.0/1.0 point (graded)

Here are a couple of definitions; the name of these techniques have been withheld:

1. A technique that divides all samples into k-groups, ideally each having all samples except one.
2. A technique that divides all samples into k-groups, ideally of equal size.
3. A technique that divides all samples into k-groups, ideally each group having the same proportion of target classes.
4. A technique that divides all samples into k-groups, ideally the same target never appearing in both testing and training groups simultaneously.

Pick the option below corresponding to the ordered definitions above:

☐ Label K-Fold, Leave-One-Out, Stratified K-Fold, K-Fold

☐ Leave-One-Out, K-Fold, Label K-Fold, Stratified K-Fold

☐ Leave-One-Out, Stratified K-Fold, K-Fold, Label K-Fold

☒ Leave-One-Out, K-Fold, Stratified K-Fold, Label K-Fold



☐ Stratified K-Fold, Leave-One-Out, K-Fold, Label K-Fold

Submit

You have used 1 of 2 attempts

## m7q6

1.0/1.0 point (graded)

There are four teaching assistants in your sister's data science course. Three of them are very sharp and completely understand the material. But one of them, unfortunately, isn't as good with machine learning as the others!

Your sister recorded the following four statements in her notebook, one made by each TA. Can you find the statement made by problematic teaching assistant?

☐ SciKit-Learn provides you with more than one powerful tool for conducting searches over your estimators' parameter values.

☐ Just about all methods called on your GridSearchCV and RandomizedSearchCV classes are actually executed against the highest scoring estimator

☒ When you build a GridSearchCV object, be sure to pass in `n_iter` so you can control the number of parameter settings that get sampled.



☐ When you pass estimator parameters into GridSearchCV, they should be in the form of an array of dictionaries. When you pass them into RandomizedSearchCV, they should be in the form of a single dictionary.

Submit

You have used 2 of 2 attempts

m7q7

0.0/1.0 point (graded)

Confusion matrices help you calculate (check all that apply):

☐ How many predictions of a target your model guessed

☒ The number of true positives + false negatives

☒ The number of false negatives + false positives

☐ How many observations per target exist in your dataset



Submit

You have used 2 of 2 attempts

© All Rights Reserved

