

Practice Quiz

Practice Quiz, 4 questions

✓ **Congratulations! You passed!**

Next Item



1 / 1
point

1.

Which of these methods are used by organizers to prevent overfitting to the test set using scores from the leaderboard?



Splitting test for public/private

Correct

The submissions are scored on public part of the test set while private score is only revealed when competition ends, thus it is not possible to overfit to private part



Shuffling order of rows in train and test sets

Un-selected is correct



Adding non-relevant samples to test set

Correct

Non-relevant samples make it much harder to do "leaderboard probing", i.e. get labels of individual objects based on leaderboard score



Limiting number of submissions per day

Correct

This also makes "leaderboard probing" much harder



1 / 1
point

2. Practice Quiz

What are the main reasons for participating in data science competitions? Check all that apply.
Practice Quiz, 4 questions

☒

To have an opportunity for networking with other participants



Correct

Yes, the people on forums are very active and helpful.

☐

To easily earn a lot of money



Un-selected is correct

☒

To learn basics of data science



Correct

Yes! For example there are plenty of public datasets on [kaggle.com](https://www.kaggle.com) with a huge community of people who is also learning.

☐

To learn things about deployment/testing



Un-selected is correct



1 / 1
point

3.

Imagine that you are participating in an image recognition competition and you want to use some publicly available set of images. Are you allowed to do this?

☐

Yes, of course

☐

No, this is not the data provided by the organizers

☒

You need to check the competition rules



Correct

If you have any doubt about competition mechanics and limitations -- check the rules or ask organizers on forums

Practice Quiz ^{1/1} point

Practice Quiz, 4 questions

4.

Which of these things we should take care about during competitions? Check all that apply.

☐

Inference speed



Un-selected is correct

☐

Deployment issues



Un-selected is correct

☒

Target metric value



Correct

This one is used for leaderboard ranking

☐

Task formalization



Un-selected is correct

