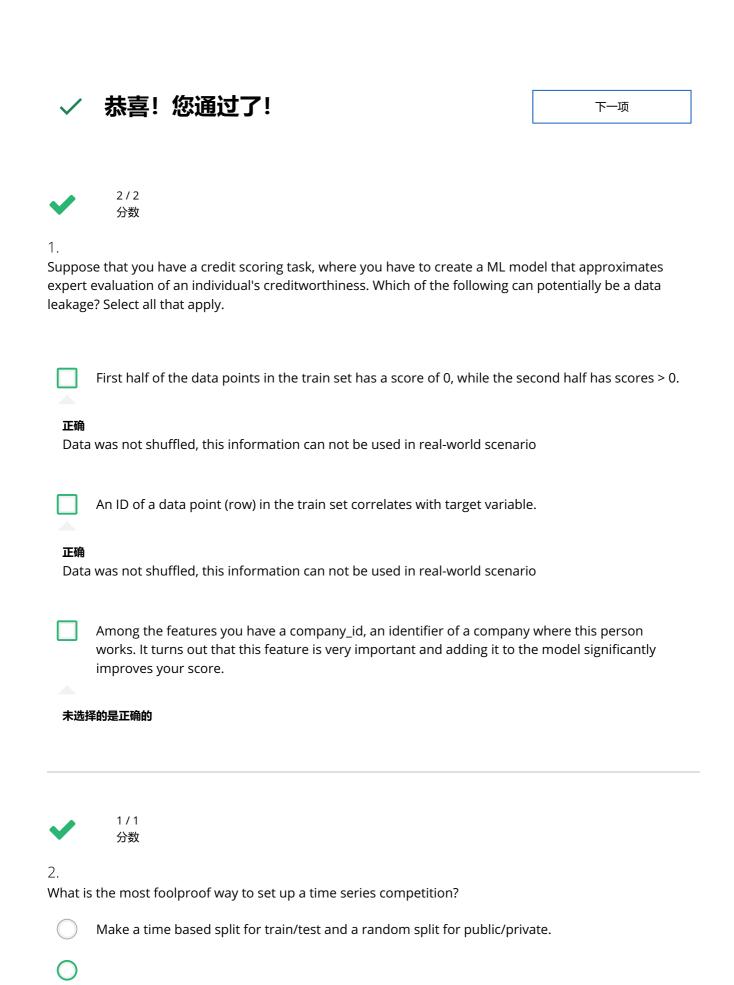
Data leakages

测验, 4 个问题

6/6 分 (100%)



Split train, public and private parts of data by time. Remove all features except IDs (e.g. Data leakagasmp) from test set so that participants will generate all the features based on past and (100%) 测验, 4 个问题 join them themselves.



正确

Correct!

Split train, public and private parts of data by time. Remove time variable from test set, keep the features.



2/2 分数

3.

Suppose that you have a binary classification task being evaluated by logloss metric. You know that there are 10000 rows in public chunk of test set and that constant 0.3 prediction gives the public score of 1.01. Mean of target variable in train is 0.44. What is the mean of target variable in public part of test data (up to 4 decimal places)?

0.7712



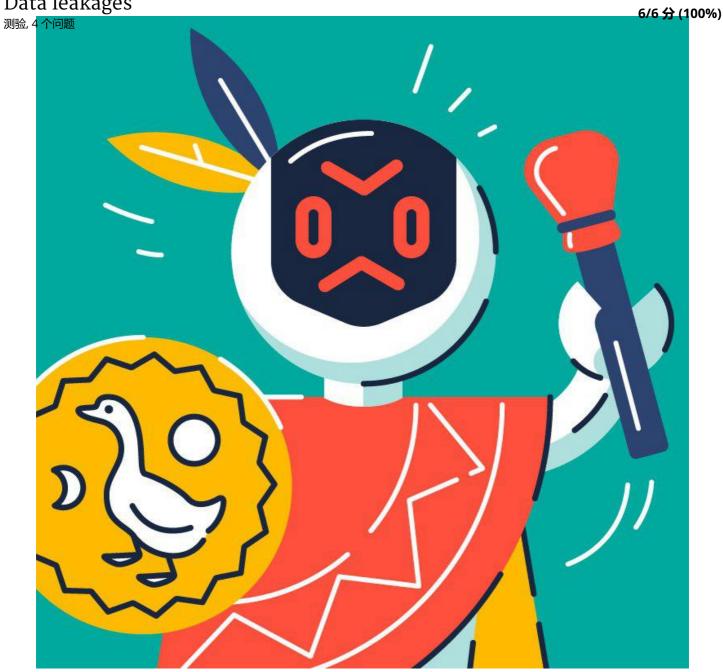


1/1

分数

4

Suppose that you are solving image classification task. What is the label of this picture? $Data\ leakages$



3

正确回答



6/6 分 (100%)