Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

Retake the assignment in **7h 56m**

GRADE
100%

Test Your Project Understanding

LATEST SUBMISSION GRADE
100%

1. Azure Machine Learning Studio experiments require no sign-in. They can be run using guest accounts. **1 / 1 point**

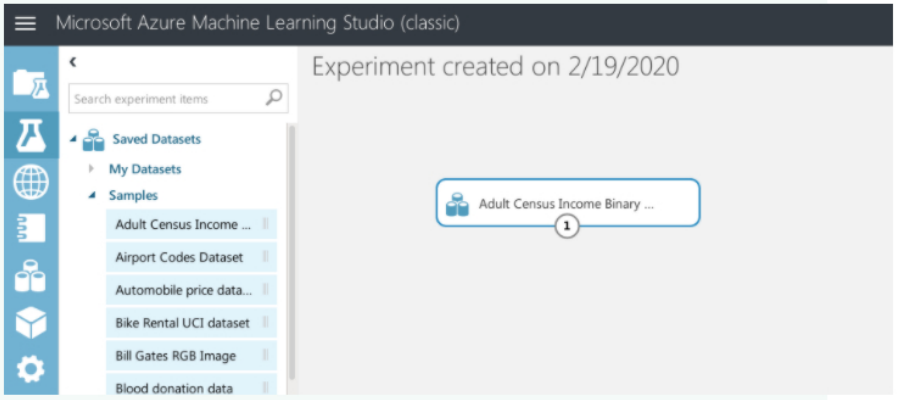
- ○ True
- ◉ False

Correct
Correct! In the reading for this hands-on project, you were guided with instructions on how to set up your Azure Machine Learning account with $200 worth of free credit to get started with running your experiments!

2. We used the Adult Census Income dataset to predict whether a household's income exceeds $50000 per year based on census data. How did we import the dataset into the blank experiment? **1 / 1 point**

- ○ We retrieved the datasets from remote URLs and imported it as modules.
- ○ We created new Azure ML datasets by uploading a local .csv file containing the data
- ◉ We used a sample dataset from Azure ML Studio

Correct
Correct! You navigated to the Saved Datasets section and selected the Adult Census Income Data from the Samples.



3. What class of machine problems does the following example belong to? **1 / 1 point**

You want to predict whether annual household income is greater than or lesser than $50000/year. The target is binary variable.

- ◉ Classification
- ○ Regression
- ○ Clustering

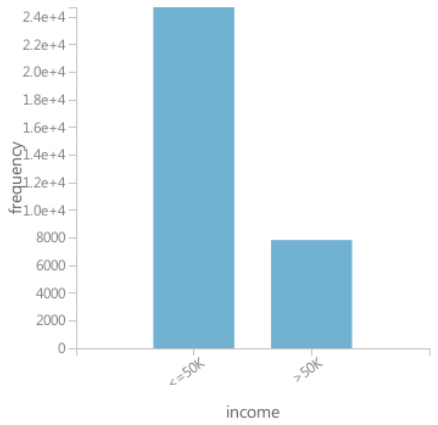
Correct
Correct!

Statistics

Unique Values	2
Missing Values	0
Feature Type	String Feature

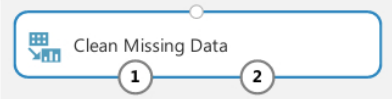
Visualizations

income
Histogram

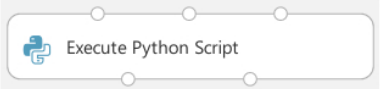


4. There are a lot of missing values in the Adult Census Income dataset. How did we deal with the missing values? **1 / 1 point**

- (●) We used the Clean Missing Data Module to subsitute all missing values with 0.



- ( ) We wrote a custom Python script to perform mean imputation, where missing values were replaced by the mean value of its corresponding column.

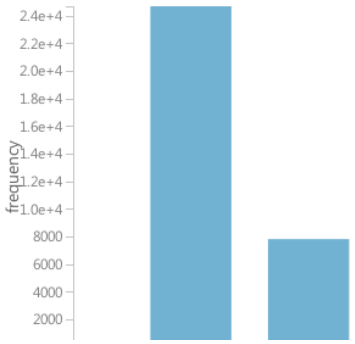


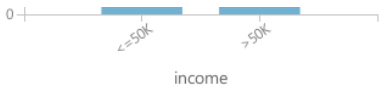
Correct
Correct!

5. The Adult Census Income dataset has a class imbalance problem. How did we overcome this challenge? **1 / 1 point**

Visualizations

income
Histogram





- ○ Oversampling (upsampling) the minority class
- ◉ Undersampling (downsampling) the majority class

Correct

Correct! We used the SMOTE module in Azure Machine Learning Studio to increase the number of underepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

6. It is best practice to perform oversampling of the minority class on the entire dataset before creating training and validation splits. **1 / 1 point**

- ○ True
- ◉ False

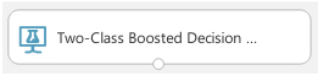
Correct

Correct! It is very easy to fall into the trap of oversampling the minority class on your entire dataset. I strongly caution against it. The timing of oversampling can affect the generalization ability of a model. Since one of the primary goals of model validation is to estimate how the model will perform on unseen data, oversampling correctly is critical. The right way to oversample is to only do it on the training data.

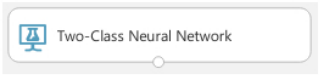
By oversampling only on the training data, none of the information in the validation data is being used to create synthetic observations. So these results should be generalizable.

7. Which of the following machine learning algorithms did we use to predict the target? **1 / 1 point**

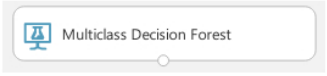
- ◉ Two-Class Boosted Decision Tree Classififer



- ○ Two-Class Neural Network



- ○ Multiclass Decision Forest



Correct

Correct!

8. Once we trained and evaluated the model, how many custom Python scripts did we write to visualize the performance of the model? **1 / 1 point**

- ◉ None of the above
- ○ 1 (a single script containing logic for both metrics)
- ○ 2 (one for each metric)

Correct

Correct! We did not have to write a single line of code! We simply right-clicked the Evaluate Model module followed by Visualize. This displayed the ROC curve, precision/recall curve, and the lift.

9. Once you have trained and scored your model, you are now ready to create a web service from an Azure Machine Learning prediction model. When the experiment run completes successfully, you will be guided to create a Scoring or Prediction Experiment. What steps are involved in preparation for deployment? (Select all that apply)

1 / 1 point

- [x] Remove one of the models

Correct

Correct!

- [x] Convert the *training experiment* you've created into a *predictive experiment*

Correct

Correct! The prediction experiment will automatically be created for you with a click. In the prediction experiment, the learner will be replaced with a trained model that has been automatically saved for you from your training experiment.

- [x] Deploy the predictive experiment as a web service

Correct

Correct! Once your scoring experiment runs successfully, you will be guided to publish your trained model as a web service.

- [ ] Set up a web server to receive test data for inference