



DP-100^{Q&As}

Designing and Implementing a Data Science Solution on Azure

Pass Microsoft DP-100 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.pass4itsure.com/dp-100.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft
Official Exam Center

- ⚙ **Instant Download** After Purchase
- ⚙ **100% Money Back** Guarantee
- ⚙ **365 Days** Free Update
- ⚙ **800,000+** Satisfied Customers



**QUESTION 1**

You are evaluating a completed binary classification machine learning model.

You need to use the precision as the evaluation metric.

Which visualization should you use?

- A. violin plot
- B. Gradient descent
- C. Scatter plot
- D. Receiver Operating Characteristic (ROC) curve

Correct Answer: D

Receiver operating characteristic (or ROC) is a plot of the correctly classified labels vs. the incorrectly classified labels for a particular model. Incorrect Answers:

A: A violin plot is a visual that traditionally combines a box plot and a kernel density plot.

B: Gradient descent is a first-order iterative optimization algorithm for finding the minimum of a function. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or approximate gradient) of the function at the current point.

C: A scatter plot graphs the actual values in your data against the values predicted by the model. The scatter plot displays the actual values along the X-axis, and displays the predicted values along the Y-axis. It also displays a line that illustrates the perfect prediction, where the predicted value exactly matches the actual value.

References: <https://docs.microsoft.com/en-us/azure/machine-learning/how-to-understand-automated-ml#confusion-matrix>

QUESTION 2

You need to define a modeling strategy for ad response.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:

**Action**

Implement a K-Means Clustering model.

Use the raw score as a feature in a Score Matchbox Recommender model.

Use the cluster as a feature in a Decision Jungle model.

Use the raw score as a feature in a Logistic Regression model.

Implement a Sweep Clustering model.

Answer area

Correct Answer:

Action

Use the raw score as a feature in a Logistic Regression model.

Implement a Sweep Clustering model.

Answer area

Implement a K-Means Clustering model.

Use the cluster as a feature in a Decision Jungle model.



Use the raw score as a feature in a Score Matchbox Recommender model.



Step 1: Implement a K-Means Clustering model

Step 2: Use the cluster as a feature in a Decision jungle model.

Decision jungles are non-parametric models, which can represent non-linear decision boundaries.

Step 3: Use the raw score as a feature in a Score Matchbox Recommender model

The goal of creating a recommendation system is to recommend one or more "items" to "users" of the system. Examples of an item could be a movie, restaurant, book, or song. A user could be a person, group of persons, or other entity with

item preferences.

Scenario:

Ad response rated declined.



Ad response models must be trained at the beginning of each event and applied during the sporting event.

Market segmentation models must optimize for similar ad response history.

Ad response models must support non-linear boundaries of features.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/multiclass-decision-jungle>

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/score-matchbox-recommender>

QUESTION 3

You plan to use a Python script to run an Azure Machine Learning experiment. The script creates a reference to the experiment run context, loads data from a file, identifies the set of unique values for the label column, and completes the experiment run:

```
from azureml.core import Runimport pandas as pdrun = Run.get_context()data = pd.read_csv('\\data.csv\\')label_vals = data[['label\\']].unique()# Add code to record metrics here run.complete()
```

The experiment must record the unique labels in the data as metrics for the run that can be reviewed later.

You must add code to the script to record the unique label values as run metrics at the point indicated by the comment.

Solution: Replace the comment with the following code:

```
run.upload_file('\\outputs/labels.csv\\', '\\./data.csv\\')
```

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

label_vals has the unique labels (from the statement `label_vals = data[['label\\']].unique()`), and it has to be logged.

Note:

Instead use the `run_log` function to log the contents in `label_vals`:

```
for label_val in label_vals: run.log('\\Label Values\\', label_val)
```

Reference: <https://www.element61.be/en/resource/azure-machine-learning-services-complete-toolbox-ai>

QUESTION 4

HOTSPOT

You create an Azure Machine Learning workspace named `workspace1`. You assign a custom role to a user of `workspace1`.



The custom role has the following JSON definition:

```
{
  "Name": "MyRole",
  "IsCustom": true,
  "Description": "New custom role description.",
  "Actions": ["*"],
  "NotActions": [
    "Microsoft.MachineLearningServices/workspaces/write",
    "Microsoft.MachineLearningServices/workspaces/computes/*/write",
    "Microsoft.MachineLearningServices/workspaces/computes/*/delete",
    "Microsoft.Authorization/*/write"
  ],
  "AssignableScopes": [
    "/subscriptions/<subscription_id>/resourceGroups/resourcegroup1/providers/
    Microsoft.MachineLearningServices/workspaces/workspaces1"
  ]
}
```

Instructions: For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

Statements	Yes	No
The user can perform all actions in the workspace	<input type="radio"/>	<input type="radio"/>
The user can delete a compute resource in the workspace	<input type="radio"/>	<input type="radio"/>
The user can write metrics to the workspace	<input type="radio"/>	<input type="radio"/>

Correct Answer:



Answer Area

Statements	Yes	No
The user can perform all actions in the workspace	<input type="radio"/>	<input checked="" type="radio"/>
The user can delete a compute resource in the workspace	<input type="radio"/>	<input checked="" type="radio"/>
The user can write metrics to the workspace	<input checked="" type="radio"/>	<input type="radio"/>

Box 1: No

The actions listed in NotActions are prohibited.

If the roles include Actions that have a wildcard (*), the effective permissions are computed by subtracting the NotActions from the allowed Actions.

Box 2: No

Deleting compute resources in the workspace is in the NotActions list.

Box 3: Yes

Writing metrics is not listed in NotActions.

Reference:

<https://docs.microsoft.com/en-us/azure/role-based-access-control/overview#how-azure-rbac-determines-if-a-user-has-access-to-a-resource>

QUESTION 5

You are a data scientist working for a bank and have used Azure ML to train and register a machine learning model that predicts whether a customer is likely to repay a loan.

You want to understand how your model is making selections and must be sure that the model does not violate government regulations such as denying loans based on where an applicant lives.

You need to determine the extent to which each feature in the customer data is influencing predictions.

What should you do?

A. Enable data drift monitoring for the model and its training dataset.

B. Score the model against some test data with known label values and use the results to calculate a confusion matrix.



- C. Use the Hyperdrive library to test the model with multiple hyperparameter values.
- D. Use the interpretability package to generate an explainer for the model.
- E. Add tags to the model registration indicating the names of the features in the training dataset.

Correct Answer: D

When you compute model explanations and visualize them, you're not limited to an existing model explanation for an automated ML model. You can also get an explanation for your model with different test data. The steps in this section show you how to compute and visualize engineered feature importance based on your test data.

Incorrect Answers:

A: In the context of machine learning, data drift is the change in model input data that leads to model performance degradation. It is one of the top reasons where model accuracy degrades over time, thus monitoring data drift helps detect model performance issues.

B: A confusion matrix is used to describe the performance of a classification model. Each row displays the instances of the true, or actual class in your dataset, and each column represents the instances of the class that was predicted by the model.

C: Hyperparameters are adjustable parameters you choose for model training that guide the training process. The HyperDrive package helps you automate choosing these parameters.

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-interpretability-automl>

[DP-100 Practice Test](#)

[DP-100 Study Guide](#)

[DP-100 Braindumps](#)