Databricks
DATABRICKS-MACHINE-LEARNING-PROFESSIONAL   Exam

Databricks Certified Machine Learning Professional

Questions & Answers
Demo

Version: 4.0

Question: 1

Which of the following describes concept drift?

A. Concept drift is when there is a change in the distribution of an input variable
B. Concept drift is when there is a change in the distribution of a target variable
C. Concept drift is when there is a change in the relationship between input variables and target variables
D. Concept drift is when there is a change in the distribution of the predicted target given by the model
E. None of these describe Concept drift

Answer: D

Question: 2

A machine learning engineer is monitoring categorical input variables for a production machine learning application. The engineer believes that missing values are becoming more prevalent in more recent data for a particular value in one of the categorical input variables.
Which of the following tools can the machine learning engineer use to assess their theory?

A. Kolmogorov-Smirnov (KS) test
B. One-way Chi-squared Test
C. Two-way Chi-squared Test
D. Jenson-Shannon distance
E. None of these

Answer: B

Question: 3

A data scientist is using MLflow to track their machine learning experiment. As a part of each MLflow run, they are performing hyperparameter tuning. The data scientist would like to have one parent run for the tuning process with a child run for each unique combination of hyperparameter values.

They are using the following code block:

```
with mlflow.start_run(run_name="Parent run") as run
    print("Start parent run")
    with mlflow.start_run(run_name="Child 1", nested=Tr
        mlflow.log_param("run_name", "child_1")
    with mlflow.start_run(run_name="Child 2", nested=Tr
        mlflow.log_param("run_name", "child_2")
```

The code block is not nesting the runs in MLflow as they expected.

Which of the following changes does the data scientist need to make to the above code block so that it successfully nests the child runs under the parent run in MLflow?

A. Indent the child run blocks within the parent run block
B. Add the nested=True argument to the parent run
C. Remove the nested=True argument from the child runs
D. Provide the same name to the run name parameter for all three run blocks
E. Add the nested=True argument to the parent run and remove the nested=True arguments from the child runs

Answer: E

Question: 4

A machine learning engineer wants to log feature importance data from a CSV file at path importance_path with an MLflow run for model model.
Which of the following code blocks will accomplish this task inside of an existing MLflow run block?
A.

```
mlflow.log_model_and_data(
    model,
    importance_path,
    "feature-importance.csv"
)
```

B.

```
mlflow.log_model(
    model,
    importance_path,
    "feature-importance.csv"
)
```

C. mlflow.log_data(importance_path, "feature-importance.csv")
D. mlflow.log_artifact(importance_path, "feature-importance.csv")
E. None of these code blocks tan accomplish the task.

Answer: A

Question: 5

Which of the following is a simple, low-cost method of monitoring numeric feature drift?

A. Jensen-Shannon test
B. Summary statistics trends
C. Chi-squared test
D. None of these can be used to monitor feature drift
E. Kolmogorov-Smirnov (KS) test

Answer: B