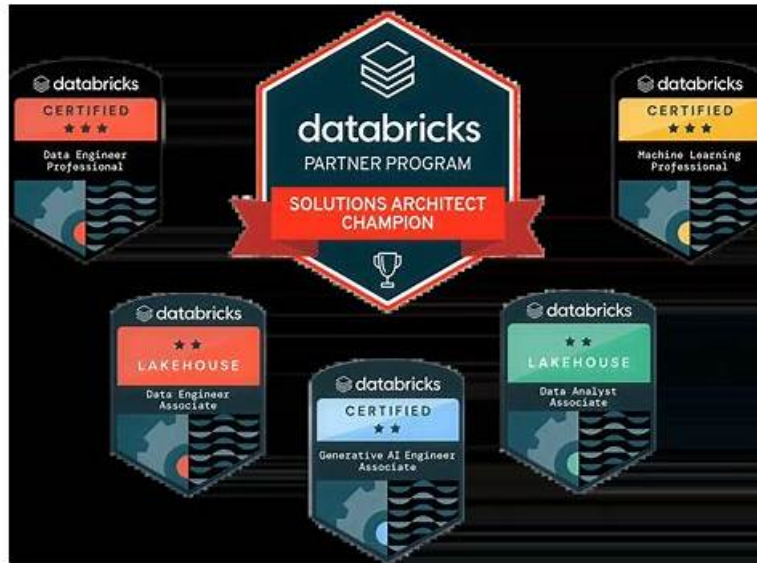


Databricks-Machine-Learning-Professional시험유형최신 기출문제공부하기



Fast2test의Databricks인증 Databricks-Machine-Learning-Professional시험대비 덤프는 가격이 착한데 비하면 품질이 너무 좋은 시험전 공부자료입니다. 시험문제적중율이 높아 패스율이 100%에 이르고 있습니다.다른 IT자격증에 관심이 있는 분들은 온라인서비스에 문의하여 덤프유무와 적중율등을 확인할수 있습니다. Databricks인증 Databricks-Machine-Learning-Professional덤프로 어려운 시험을 정복하여 IT업계 정상에 오릅니다.

Databricks Databricks-Machine-Learning-Professional 시험요강:

주제	소개
주제 1	- Identify a use case for HTTP webhooks and where the Webhook URL needs to come - Identify advantages of using Job clusters over all-purpose clusters
주제 2	- Identify less performant data storage as a solution for other use cases - Describe why complex business logic must be handled in streaming deployments
주제 3	- Describe concept drift and its impact on model efficacy - Describe summary statistic monitoring as a simple solution for numeric feature drift
주제 4	- Identify that data can arrive out-of-order with structured streaming - Identify how model serving uses one all-purpose cluster for a model deployment
주제 5	- Identify live serving benefits of querying precomputed batch predictions - Describe Structured Streaming as a common processing tool for ETL pipelines
주제 6	- Test whether the updated model performs better on the more recent data - Identify when retraining and deploying an updated model is a probable solution to drift
주제 7	- Create, overwrite, merge, and read Feature Store tables in machine learning workflows - View Delta table history and load a previous version of a Delta table
주제 8	- Describe the advantages of using the pyfunc MLflow flavor - Manually log parameters, models, and evaluation metrics using MLflow

Databricks Databricks-Machine-Learning-Professional최고덤프 - Databricks-Machine-Learning-Professional완벽한 공부문제

Databricks Databricks-Machine-Learning-Professional덤프의 무료샘플을 원하신다면 우의 PDF Version Demo 버튼을 클릭하고 메일주소를 입력하시면 바로 다운받아Databricks Databricks-Machine-Learning-Professional덤프의 일부분 문제를 체험해 보실수 있습니다. Databricks Databricks-Machine-Learning-Professional 덤프는 모든 시험문제유형을 포함하고 있어 적응율이 아주 높습니다. Databricks Databricks-Machine-Learning-Professional덤프로Databricks Databricks-Machine-Learning-Professional시험패스 GO GO GO !

최신 ML Data Scientist Databricks-Machine-Learning-Professional 무료샘플문제 (Q179-Q184):

질문 # 179

A Machine Learning Engineer has automated a model retraining job in Databricks. Each scheduled run trains multiple candidate models with new sales data and logs all runs with MLflow.

The goal is to select and register the best-performing model at the end of each cycle to ensure optimal forecast accuracy. Which approach will meet this goal?

- A. Register all models from the run since all of them are trained on newer data and hence will be more accurate.
- B. Register the model with the lowest training loss regardless of validation metrics.
- C. Register the best model based on the primary evaluation metric.
- D. Register the first model that has an evaluation metric better than the previously deployed model.

정답: C

설명:

Selecting and registering the model that performs best on the primary evaluation metric ensures that only the highest-quality model is promoted at each retraining cycle. This approach aligns with MLOps best practices by basing promotion decisions on objective performance criteria rather than training order or assumptions about data freshness.

질문 # 180

A machine learning engineer has registered a sklearn model in the MLflow Model Registry using the sklearn model flavor with UI model_uri.

Which of the following operations can be used to load the model as an sklearn object for batch deployment?

- A. mlflow.pyfunc.read_model(model_uri)
- B. mlflow.pyfunc.load_model(model_uri)
- C. mlflow.sklearn.load_model(model_uri)
- D. mlflow.spark.load_model(model_uri)
- E. mlflow.sklearn.read_model(model_uri)

정답: B

질문 # 181

A Machine Learning Engineer is using Lakehouse Monitoring to track the performance of ML models deployed in their environment. They want to monitor significant distributional drift in categorical features with a metric bounded on $[0,1]$ for easy interpretation.

Which statistical method should they use?

- A. Kolmogorov-Smirnov Test
- B. Kullback-Leibler Distance
- C. Jensen-Shannon Distance
- D. Wasserstein Distance

정답: C

설명:

Jensen-Shannon Distance is well suited for measuring distributional drift in categorical features. It is symmetric, numerically stable, and bounded between 0 and 1, which makes it easy to interpret and ideal for monitoring categorical feature drift in Lakehouse

Monitoring.

질문 # 182

A machine learning engineer has deployed a model recommender using MLflow Model Serving.

They now want to query the version of that model that is in the Staging stage of the MLflow Model Registry. Which model URI can be used to query the described model version?

- A. The version number of the model version in Production is necessary to complete this task.
- B. <https://<databricks-instance>/model/recommender/Staging/invocations>
- C. <https://<databricks-instance>/model-serving/recommender/Staging/invocations>
- D. It is not possible to query a model version in the Staging stage.

정답: B

설명:

To query a model version in a specific stage (like Staging) using MLflow Model Serving on Databricks, the correct model URI format is:

<https://<databricks-instance>/model/<model-name>/<stage>/invocations>

So, for the model named recommender in the Staging stage, the correct URI is:

<https://<databricks-instance>/model/recommender/Staging/invocations>.

질문 # 183

A Machine Learning Engineer has deployed a fraud detection model that processes 10,000 transactions per hour. The model was trained on data from Q1 2024, but it's now Q4 2024. The ML team notices three concerning trends: (1) the model's precision has dropped from 92% to

78% over the past month, (2) the average transaction amount in recent data has increased from

\$150 to \$220 and (3) the relationship between transaction frequency and fraud likelihood has weakened significantly due to new payment methods being introduced. The engineer needs to implement a monitoring solution that can detect the root cause of the performance degradation, identify why the precision dropped, and be able to do this at the scale needed. Which monitoring pipeline component will do this?

- A. Logging and Monitoring Pipeline because it captures all prediction requests and can identify patterns in the fraud predictions.
- B. Model Health Monitoring Pipeline because it ensures the infrastructure can handle the 10,000 transactions per hour load.
- C. Model Performance Monitoring Pipeline because it directly tracks precision metrics and can alert when thresholds are breached.
- D. [Drift Detection Pipeline because it can identify both the data drift \(transaction amount changes\) and concept drift \(weakened relationship between features and target\).](#)

정답: D

설명:

A drift detection pipeline is designed to diagnose the root causes of model performance degradation at scale. It can detect data drift, such as changes in the distribution of transaction amounts, and concept drift, where the relationship between input features and the fraud label changes due to new payment methods. This directly explains why precision dropped and provides actionable insight beyond simply observing metric degradation.

질문 # 184

.....

경쟁율이 점점 높아지는 IT업계에 살아남으려면 국제적으로 인증해주는 IT자격증 몇개쯤은 취득해야 되지 않을까요? Databricks Databricks-Machine-Learning-Professional 시험으로부터 자격증 취득을 시작해보세요. Databricks Databricks-Machine-Learning-Professional 덤프의 모든 문제를 외우기만 하면 시험패스가 됩니다. Databricks Databricks-Machine-Learning-Professional 덤프는 실제 시험문제의 모든 유형을 포함되어 있어 적응율이 최고입니다.

Databricks-Machine-Learning-Professional 최고 덤프 : <https://kr.fast2test.com/Databricks-Machine-Learning-Professional-premium-file.html>

- [illegible]