MLS-C01 Q&As

AWS Certified Machine Learning - Specialty (MLS-C01)

Pass Amazon MLS-C01 Exam with 100% Guarantee

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

https://www.geekcert.com/aws-certified-machine-learning-specialty.html

100% Passing Guarantee
100% Money Back Assurance

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

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

QUESTION 1

A data scientist is working on a public sector project for an urban traffic system. While studying the traffic patterns, it is clear to the data scientist that the traffic behavior at each light is correlated, subject to a small stochastic error term. The data scientist must model the traffic behavior to analyze the traffic patterns and reduce congestion.

How will the data scientist MOST effectively model the problem?

A. The data scientist should obtain a correlated equilibrium policy by formulating this problem as a multi-agent reinforcement learning problem.

B. The data scientist should obtain the optimal equilibrium policy by formulating this problem as a single-agent reinforcement learning problem.

C. Rather than finding an equilibrium policy, the data scientist should obtain accurate predictors of traffic flow by using historical data through a supervised learning approach.

D. Rather than finding an equilibrium policy, the data scientist should obtain accurate predictors of traffic flow by using unlabeled simulated data representing the new traffic patterns in the city and applying an unsupervised learning approach.

Correct Answer: A

Reference: https://www.hindawi.com/journals/jat/2021/8878011/

QUESTION 2

A company wants to create an artificial intelligence (AI) yoga instructor that can lead large classes of students. The company needs to create a feature that can accurately count the number of students who are in a class. The company also needs a feature that can differentiate students who are performing a yoga stretch correctly from students who are performing a stretch incorrectly.

...etermine whether students are performing a stretch correctly, the solution needs to measure the location and angle of each student\\'s arms and legs A data scientist must use Amazon SageMaker to ...ss video footage of a yoga class by extracting image frames and applying computer vision models.

Which combination of models will meet these requirements with the LEAST effort? (Select TWO.)

A. Image Classification

B. Optical Character Recognition (OCR)

C. Object Detection

D. Pose estimation

E. Image Generative Adversarial Networks (GANs)

Correct Answer: CD

To count the number of students who are in a class, the solution needs to detect and locate each student in the video frame. Object detection is a computer vision model that can identify and locate multiple objects in an image. To differentiate

students who are performing a stretch correctly from students who are performing a stretch incorrectly, the solution needs to measure the location and angle of each student\\'s arms and legs. Pose estimation is a computer vision model that

can estimate the pose of a person by detecting the position and orientation of key body parts. Image classification, OCR, and image GANs are not relevant for this use case.

References:

Object Detection: A computer vision technique that identifies and locates objects within an image or video. Pose Estimation: A computer vision technique that estimates the pose of a person by detecting the position and orientation of key body

parts. Amazon SageMaker: A fully managed service that provides every developer and data scientist with the ability to build, train, and deploy machine learning (ML) models quickly.

QUESTION 3

A data engineer wants to perform exploratory data analysis (EDA) on a petabyte of data. The data engineer does not want to manage compute resources and wants to pay only for queries that are run. The data engineer must write the analysis by using Python from a Jupyter notebook.

Which solution will meet these requirements?

A. Use Apache Spark from within Amazon Athena.

B. Use Apache Spark from within Amazon SageMaker.

C. Use Apache Spark from within an Amazon EMR cluster.

D. Use Apache Spark through an integration with Amazon Redshift.

Correct Answer: B

QUESTION 4

A machine learning (ML) specialist is using the Amazon SageMaker DeepAR forecasting algorithm to train a model on CPU-based Amazon EC2 On-Demand instances. The model currently takes multiple hours to train. The ML specialist wants to decrease the training time of the model.

Which approaches will meet this requirement? (SELECT TWO)

A. Replace On-Demand Instances with Spot Instances

B. Configure model auto scaling dynamically to adjust the number of instances automatically.

C. Replace CPU-based EC2 instances with GPU-based EC2 instances.

D. Use multiple training instances.

E. Use a pre-trained version of the model. Run incremental training.

Correct Answer: CD

The best approaches to decrease the training time of the model are C and D, because they can improve the computational efficiency and parallelization of the training process. These approaches have the following benefits:

C: Replacing CPU-based EC2 instances with GPU-based EC2 instances can speed up the training of the DeepAR algorithm, as it can leverage the parallel processing power of GPUs to perform matrix operations and gradient computations

faster than CPUs12. The DeepAR algorithm supports GPU-based EC2 instances such as ml.p2 and ml.p33.

D: Using multiple training instances can also reduce the training time of the DeepAR algorithm, as it can distribute the workload across multiple nodes and perform data parallelism4. The DeepAR algorithm supports distributed training with

multiple CPU-based or GPU-based EC2 instances3. The other options are not effective or relevant, because they have the following drawbacks:

A: Replacing On-Demand Instances with Spot Instances can reduce the cost of the training, but not necessarily the time, as Spot Instances are subject to interruption and availability5. Moreover, the DeepAR algorithm does not support

checkpointing, which means that the training cannot resume from the last saved state if the Spot Instance is terminated3.

B: Configuring model auto scaling dynamically to adjust the number of instances automatically is not applicable, as this feature is only available for inference endpoints, not for training jobs6.

E: Using a pre-trained version of the model and running incremental training is not possible, as the DeepAR algorithm does not support incremental training or transfer learning3. The DeepAR algorithm requires a full retraining of the model

whenever new data is added or the hyperparameters are changed7.

References:

1: GPU vs CPU: What Matters Most for Machine Learning? | by Louis (What\\'s AI) Bouchard | Towards Data Science

2: How GPUs Accelerate Machine Learning Training | NVIDIA Developer Blog

3: DeepAR Forecasting Algorithm - Amazon SageMaker

4: Distributed Training - Amazon SageMaker

5: Managed Spot Training - Amazon SageMaker

6: Automatic Scaling - Amazon SageMaker

7: How the DeepAR Algorithm Works - Amazon SageMaker

QUESTION 5

A data scientist at a food production company wants to use an Amazon SageMaker built-in model to classify different vegetables. The current dataset has many features. The company wants to save on memory costs when the data scientist trains and deploys the model. The company also wants to be able to find similar data points for each test data point.

Which algorithm will meet these requirements?

A. K-nearest neighbors (k-NN) with dimension reduction

B. Linear learner with early stopping

C. K-means

D. Principal component analysis (PCA) with the algorithm mode set to random

Correct Answer: A