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.pass4itsure.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 company is running a machine learning prediction service that generates 100 TB of predictions every day A Machine Learning Specialist must generate a visualization of the daily precision-recall curve from the predictions, and forward a read-only version to the Business team.

Which solution requires the LEAST coding effort?

A. Run a daily Amazon EMR workflow to generate precision-recall data, and save the results in Amazon S3 Give the Business team read-only access to S3

B. Generate daily precision-recall data in Amazon QuickSight, and publish the results in a dashboard shared with the Business team

C. Run a daily Amazon EMR workflow to generate precision-recall data, and save the results in Amazon

D. S3 Visualize the arrays in Amazon QuickSight, and publish them in a dashboard shared with the Business team

E. Generate daily precision-recall data in Amazon ES, and publish the results in a dashboard shared with the Business team.

Correct Answer: C

QUESTION 2

An ecommerce company is automating the categorization of its products based on images. A data scientist has trained a computer vision model using the Amazon SageMaker image classification algorithm. The images for each product are classified according to specific product lines. The accuracy of the model is too low when categorizing new products. All of the product images have the same dimensions and are stored within an Amazon S3 bucket. The company wants to improve the model so it can be used for new products as soon as possible.

Which steps would improve the accuracy of the solution? (Choose three.)

A. Use the SageMaker semantic segmentation algorithm to train a new model to achieve improved accuracy.

B. Use the Amazon Rekognition DetectLabels API to classify the products in the dataset.

C. Augment the images in the dataset. Use open source libraries to crop, resize, flip, rotate, and adjust the brightness and contrast of the images.

D. Use a SageMaker notebook to implement the normalization of pixels and scaling of the images. Store the new dataset in Amazon S3.

E. Use Amazon Rekognition Custom Labels to train a new model.

F. Check whether there are class imbalances in the product categories, and apply oversampling or undersampling as required. Store the new dataset in Amazon S3.

Correct Answer: CEF

Reference: https://docs.aws.amazon.com/rekognition/latest/dg/how-it-works-types.html
https://towardsdatascience.com/image-processing-techniques-for-computer-vision-11f92f511e21
https://docs.aws.amazon.com/rekognition/latest/customlabels-dg/training-model.html

QUESTION 3

A network security vendor needs to ingest telemetry data from thousands of endpoints that run all over the world. The data is transmitted every 30 seconds in the form of records that contain 50 fields. Each record is up to 1 KB in size. The security vendor uses Amazon Kinesis Data Streams to ingest the data. The vendor requires hourly summaries of the records that Kinesis Data Streams ingests. The vendor will use Amazon Athena to query the records and to generate the summaries. The Athena queries will target 7 to 12 of the available data fields.

Which solution will meet these requirements with the LEAST amount of customization to transform and store the ingested data?

A. Use AWS Lambda to read and aggregate the data hourly. Transform the data and store it in Amazon S3 by using Amazon Kinesis Data Firehose.

B. Use Amazon Kinesis Data Firehose to read and aggregate the data hourly. Transform the data and store it in Amazon S3 by using a short-lived Amazon EMR cluster.

C. Use Amazon Kinesis Data Analytics to read and aggregate the data hourly. Transform the data and store it in Amazon S3 by using Amazon Kinesis Data Firehose.

D. Use Amazon Kinesis Data Firehose to read and aggregate the data hourly. Transform the data and store it in Amazon S3 by using AWS Lambda.

Correct Answer: C

QUESTION 4

A Machine Learning Specialist receives customer data for an online shopping website. The data includes demographics, past visits, and locality information. The Specialist must develop a machine learning approach to identify the customer shopping patterns, preferences and trends to enhance the website for better service and smart recommendations.

Which solution should the Specialist recommend?

A. Latent Dirichlet Allocation (LDA) for the given collection of discrete data to identify patterns in the customer database.

B. A neural network with a minimum of three layers and random initial weights to identify patterns in the customer database

C. Collaborative filtering based on user interactions and correlations to identify patterns in the customer database

D. Random Cut Forest (RCF) over random subsamples to identify patterns in the customer database

Correct Answer: C

In natural language processing, the latent Dirichlet allocation (LDA) is a generative statistical model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar.

Amazon SageMaker Random Cut Forest (RCF) is an unsupervised algorithm for detecting anomalous data points within a data set

Neural network is used for image detection.

QUESTION 5

A global financial company is using machine learning to automate its loan approval process. The company has a dataset of customer information. The dataset contains some categorical fields, such as customer location by city and housing

status. The dataset also includes financial fields in different units, such as account balances in US dollars and monthly interest in US cents.

The company\\'s data scientists are using a gradient boosting regression model to infer the credit score for each customer. The model has a training accuracy of 99% and a testing accuracy of 75%. The data scientists want to improve the model\\'s testing accuracy.

Which process will improve the testing accuracy the MOST?

A. Use a one-hot encoder for the categorical fields in the dataset. Perform standardization on the financial fields in the dataset. Apply L1 regularization to the data.

B. Use tokenization of the categorical fields in the dataset. Perform binning on the financial fields in the dataset. Remove the outliers in the data by using the z-score.

C. Use a label encoder for the categorical fields in the dataset. Perform L1 regularization on the financial fields in the dataset. Apply L2 regularization to the data.

D. Use a logarithm transformation on the categorical fields in the dataset. Perform binning on the financial fields in the dataset. Use imputation to populate missing values in the dataset.

Correct Answer: A

MLS-C01 PDF Dumps MLS-C01 Practice Test MLS-C01 Study Guide