

MLS-C01 Latest Dumps Sheet & MLS-C01 Valid Exam Prep



BTW, DOWNLOAD part of ExamBoosts MLS-C01 dumps from Cloud Storage: https://drive.google.com/open?id=1BzUlk7ixoCPQKTIJZGOv9y---tl_aqaY

The ExamBoosts is a trusted and leading platform that is committed to making the entire Amazon MLS-C01 exam preparation process simple, smart, and quick. To achieve this objective ExamBoosts is offering real, valid, and updated Amazon MLS-C01 Exam Questions. These Amazon MLS-C01 exam dumps are the real MLS-C01 exam questions that surely will repeat in the upcoming MLS-C01 exam and you can pass the challenging exam.

Amazon AWS Certified Machine Learning Specialty Exam is intended for individuals who have a strong understanding of data analytics and machine learning, as well as experience working with AWS architectures. MLS-C01 Exam requires knowledge of how to navigate AWS services and tools, including Amazon Sagemaker, Amazon Comprehend, and Amazon Rekognition. Candidates should also have experience using programming languages like Python and have a deep understanding of the fundamentals of machine learning.

[**>> MLS-C01 Latest Dumps Sheet <<**](#)

MLS-C01 Valid Exam Prep, MLS-C01 Reliable Exam Voucher

The MLS-C01 learning dumps from our company are very convenient for all people, including the convenient buying process, the download way and the study process and so on. Upon completion of your payment, you will receive the email from us in several minutes, and then you will have the right to use the AWS Certified Machine Learning - Specialty test guide from our company. In addition, there are three different versions for all people to choose. According to your actual situation, you can choose the suitable version from our MLS-C01 study question. We believe that the suitable version will help you improve your learning efficiency. It will be very easy for you to pass the exam and get the certification. More importantly, your will spend less time on preparing for MLS-C01 exam than other people.

Amazon AWS Certified Machine Learning - Specialty Sample Questions (Q170-Q175):

NEW QUESTION # 170

When submitting Amazon SageMaker training jobs using one of the built-in algorithms, which common parameters MUST be specified? (Select THREE.)

- A. Hyperparameters in a JSON array as documented for the algorithm used.
- B. The output path specifying where on an Amazon S3 bucket the trained model will persist.
- C. The Amazon EC2 instance class specifying whether training will be run using CPU or GPU.
- D. The IAM role that Amazon SageMaker can assume to perform tasks on behalf of the users.
- E. The validation channel identifying the location of validation data on an Amazon S3 bucket.
- F. The training channel identifying the location of training data on an Amazon S3 bucket.

Answer: B,D,F

Explanation:

When submitting Amazon SageMaker training jobs using one of the built-in algorithms, the common parameters that must be specified are:

The training channel identifying the location of training data on an Amazon S3 bucket. This parameter tells SageMaker where to find the input data for the algorithm and what format it is in. For example, TrainingInputMode: File means that the input data is in files stored in S3.

The IAM role that Amazon SageMaker can assume to perform tasks on behalf of the users. This parameter grants SageMaker the necessary permissions to access the S3 buckets, ECR repositories, and other AWS resources needed for the training job. For example, RoleArn: arn:aws:iam:123456789012:role/service-role

/AmazonSageMaker-ExecutionRole-20200303T150948 means that SageMaker will use the specified role to run the training job.

The output path specifying where on an Amazon S3 bucket the trained model will persist. This parameter tells SageMaker where to save the model artifacts, such as the model weights and parameters, after the training job is completed. For example, OutputDataConfig: {S3OutputPath: s3://my-bucket/my-training-job} means that SageMaker will store the model artifacts in the specified S3 location.

The validation channel identifying the location of validation data on an Amazon S3 bucket is an optional parameter that can be used to provide a separate dataset for evaluating the model performance during the training process. This parameter is not required for all algorithms and can be omitted if the validation data is not available or not needed.

The hyperparameters in a JSON array as documented for the algorithm used is another optional parameter that can be used to customize the behavior and performance of the algorithm. This parameter is specific to each algorithm and can be used to tune the model accuracy, speed, complexity, and other aspects. For example, HyperParameters: {num_round: "10", objective: "binary:logistic"} means that the XGBoost algorithm will use 10 boosting rounds and the logistic loss function for binary classification. The Amazon EC2 instance class specifying whether training will be run using CPU or GPU is not a parameter that is specified when submitting a training job using a built-in algorithm. Instead, this parameter is specified when creating a training instance, which is a containerized environment that runs the training code and algorithm. For example, ResourceConfig: {InstanceType: ml.m5.xlarge, InstanceCount: 1, VolumeSizeInGB:

10} means that SageMaker will use one m5.xlarge instance with 10 GB of storage for the training instance.

Train a Model with Amazon SageMaker

Use Amazon SageMaker Built-in Algorithms or Pre-trained Models

CreateTrainingJob - Amazon SageMaker Service

NEW QUESTION # 171

A company is launching a new product and needs to build a mechanism to monitor comments about the company and its new product on social media. The company needs to be able to evaluate the sentiment expressed in social media posts, and visualize trends and configure alarms based on various thresholds.

The company needs to implement this solution quickly, and wants to minimize the infrastructure and data science resources needed to evaluate the messages. The company already has a solution in place to collect posts and store them within an Amazon S3 bucket. What services should the data science team use to deliver this solution?

- A. Trigger an AWS Lambda function when social media posts are added to the S3 bucket. Call Amazon Comprehend for each post to capture the sentiment in the message and record the sentiment in a custom Amazon CloudWatch metric and in S3. Use CloudWatch alarms to notify analysts of trends.
- B. Train a model in Amazon SageMaker by using the BlazingText algorithm to detect sentiment in the corpus of social media posts. Expose an endpoint that can be called by AWS Lambda. Trigger a Lambda function when posts are added to the S3 bucket to invoke the endpoint and record the sentiment in an Amazon DynamoDB table and in a custom Amazon CloudWatch metric. Use CloudWatch alarms to notify analysts of trends.
- C. Train a model in Amazon SageMaker by using the semantic segmentation algorithm to model the semantic content in the corpus of social media posts. Expose an endpoint that can be called by AWS Lambda. Trigger a Lambda function when

objects are added to the S3 bucket to invoke the endpoint and record the sentiment in an Amazon DynamoDB table. Schedule a second Lambda function to query recently added records and send an Amazon Simple Notification Service (Amazon SNS) notification to notify analysts of trends.

- D. Trigger an AWS Lambda function when social media posts are added to the S3 bucket. Call Amazon Comprehend for each post to capture the sentiment in the message and record the sentiment in an Amazon DynamoDB table. Schedule a second Lambda function to query recently added records and send an Amazon Simple Notification Service (Amazon SNS) notification to notify analysts of trends.

Answer: A

Explanation:

Explanation

The solution that uses Amazon Comprehend and Amazon CloudWatch is the most suitable for the given scenario. Amazon Comprehend is a natural language processing (NLP) service that can analyze text and extract insights such as sentiment, entities, topics, and syntax. Amazon CloudWatch is a monitoring and observability service that can collect and track metrics, create dashboards, and set alarms based on various thresholds. By using these services, the data science team can quickly and easily implement a solution to monitor the sentiment of social media posts without requiring much infrastructure or data science resources. The solution also meets the requirements of storing the sentiment in both S3 and CloudWatch, and using CloudWatch alarms to notify analysts of trends.

References:

Amazon Comprehend

Amazon CloudWatch

NEW QUESTION # 172

A Machine Learning Specialist must build out a process to query a dataset on Amazon S3 using Amazon Athena. The dataset contains more than 800,000 records stored as plaintext CSV files. Each record contains 200 columns and is approximately 1.5 MB in size. Most queries will span 5 to 10 columns only. How should the Machine Learning Specialist transform the dataset to minimize query runtime?

- A. Convert the records to XML format
- B. Convert the records to JSON format
- C. Convert the records to Apache Parquet format
- D. Convert the records to GZIP CSV format

Answer: C

Explanation:

* Explanation: Amazon Athena is an interactive query service that allows you to analyze data stored in Amazon S3 using standard SQL. Athena is serverless, so you only pay for the queries that you run and there is no infrastructure to manage.

* To optimize the query performance of Athena, one of the best practices is to convert the data into a columnar format, such as Apache Parquet or Apache ORC. Columnar formats store data by columns rather than by rows, which allows Athena to scan only the columns that are relevant to the query, reducing the amount of data read and improving the query speed. Columnar formats also support compression and encoding schemes that can reduce the storage space and the data scanned per query, further enhancing the performance and reducing the cost.

* In contrast, plaintext CSV files store data by rows, which means that Athena has to scan the entire row even if only a few columns are needed for the query. This increases the amount of data read and the query latency. Moreover, plaintext CSV files do not support compression or encoding, which means that they take up more storage space and incur higher query costs.

* Therefore, the Machine Learning Specialist should transform the dataset to Apache Parquet format to minimize query runtime.

References:

* Top 10 Performance Tuning Tips for Amazon Athena

* Columnar Storage Formats

Using compressions will reduce the amount of data scanned by Amazon Athena, and also reduce your S3 bucket storage. It's a Win-Win for your AWS bill. Supported formats: GZIP, LZO, SNAPPY (Parquet) and ZLIB.

NEW QUESTION # 173

A Machine Learning Specialist is working with a media company to perform classification on popular articles from the company's website. The company is using random forests to classify how popular an article will be before it is published. A sample of the data being used is below.

Given the dataset, the Specialist wants to convert the Day-Of_Week column to binary values.

What technique should be used to convert this column to binary values.

Article_Title	Author	Top_Keywords	Day_Of_Week	URL of Article	Page_VIEWS
Building a Big Data Platform	Jane Doe	Big Data, Spark, Hadoop	Tuesday	http://examplecorp.com/data_platform.html	1300456
Getting Started with Deep Learning	John Doe	Deep Learning, Machine Learning, Spark	Tuesday	http://examplecorp.com/started_deep_learning.html	1230661
MXNet ML Guide	Jane Doe	Machine Learning, MXNet, Logistic Regression	Thursday	http://examplecorp.com/mxnet_guide.html	937291
Intro to NoSQL Databases	Mary Major	NoSQL, Operations, Database	Monday	http://examplecorp.com/nosql_intro_guide.html	407812

- A. Tokenization
- B. Binarization
- C. One-hot encoding**
- D. Normalization transformation

Answer: C

NEW QUESTION # 174

A Machine Learning Specialist is assigned to a Fraud Detection team and must tune an XGBoost model, which is working appropriately for test data. However, with unknown data, it is not working as expected. The existing parameters are provided as follows.

```
param = {
    'eta': 0.05, # the training step for each iteration
    'silent': 1, # logging mode - quiet
    'n_estimators':2000,
    'max_depth':30,
    'min_child_weight': 3,
    'gamma': 0,
    'subsample': 0.8,
    'objective': 'multi:softprob', # error evaluation for multiclass training
    'num_class': 201} # the number of classes that exist in this dataset
num_round = 60 # the number of training iterations
```

Which parameter tuning guidelines should the Specialist follow to avoid overfitting?

- A. Lower the max_depth parameter value.**
- B. Update the objective to binary:logistic.
- C. Lower the min_child_weight parameter value.
- D. Increase the max_depth parameter value.

Answer: A

NEW QUESTION # 175

.....

The price for MLS-C01 study guide is quite reasonable, no matter you are a student or employee in the company, you can afford them. Just think that, you only need to spend some money, you can get a certificate as well as improve your ability. Besides, we also pass guarantee and money back guarantee for you fail to pass the exam after you have purchasing MLS-C01 Exam Dumps from us. We can give you free update for 365 days after your purchasing. If you have any questions about the MLS-C01 study guide, you can have a chat with us.

MLS-C01 Valid Exam Prep: <https://www.examboosts.com/Amazon/MLS-C01-practice-exam-dumps.html>

DOWNLOAD the newest ExamBoosts MLS-C01 PDF dumps from Cloud Storage for free: https://drive.google.com/open?id=1BzUlk7ixoCPQKTIJZGOv9y---tl_aqaY