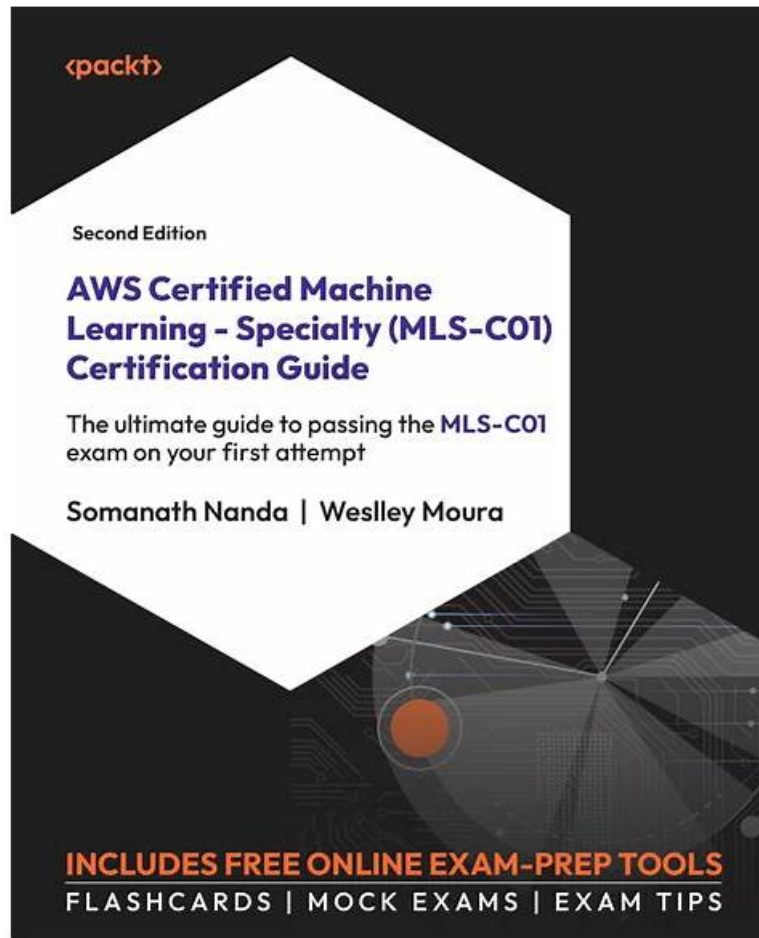


MLS-C01 Valid Learning Materials & Reliable MLS-C01 Exam Registration



BONUS!!! Download part of UpdateDumps MLS-C01 dumps for free: <https://drive.google.com/open?id=1fLdFvg8dlgBTYqlovsGaufokLy2A100>

As usual, you just need to spend little time can have a good commend of our study materials, then you can attend to your MLS-C01 exam and pass it at your first attempt. We also hire a team of experts, and the content of MLS-C01 question torrent is all high-quality test guidance materials that have been accepted by experienced professionals. MLS-C01 practice materials will be the most professional and dedicated tutor you have ever met.

The AWS Certified Machine Learning - Specialty exam is a certification offered by Amazon Web Services (AWS) for professionals who want to validate their expertise in machine learning. MLS-C01 exam is designed for individuals who have a solid understanding of machine learning concepts and techniques, as well as experience using AWS services to build and deploy machine learning solutions. MLS-C01 Exam covers a range of topics, including data preparation, model training and evaluation, deployment and implementation, and automation.

>> **MLS-C01 Valid Learning Materials** <<

Reliable Amazon MLS-C01 Exam Registration - Exam MLS-C01 Online

Candidates can reach out to the UpdateDumps support staff anytime. The UpdateDumps help desk is the place to go if you have any questions or problems. Time management is crucial to passing the Amazon MLS-C01 exam. Candidates may prepare for the Amazon MLS-C01 Exam with the help of UpdateDumps desktop-based MLS-C01 practice exam software, web-based MLS-C01 practice tests and Amazon MLS-C01 pdf questions.

Amazon AWS Certified Machine Learning - Specialty Sample Questions (Q190-Q195):

NEW QUESTION # 190

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

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

Answer: A,E,F

NEW QUESTION # 191

A Data Science team is designing a dataset repository where it will store a large amount of training data commonly used in its machine learning models. As Data Scientists may create an arbitrary number of new datasets every day the solution has to scale automatically and be cost-effective. Also, it must be possible to explore the data using SQL.

Which storage scheme is MOST adapted to this scenario?

- A. Store datasets as tables in a multi-node Amazon Redshift cluster.
- B. Store datasets as files in Amazon S3.
- C. Store datasets as global tables in Amazon DynamoDB.
- D. Store datasets as files in an Amazon EBS volume attached to an Amazon EC2 instance.

Answer: B

Explanation:

The best storage scheme for this scenario is to store datasets as files in Amazon S3. Amazon S3 is a scalable, cost-effective, and durable object storage service that can store any amount and type of data. Amazon S3 also supports querying data using SQL with Amazon Athena, a serverless interactive query service that can analyze data directly in S3. This way, the Data Science team can easily explore and analyze their datasets without having to load them into a database or a compute instance.

The other options are not as suitable for this scenario because:

* Storing datasets as files in an Amazon EBS volume attached to an Amazon EC2 instance would limit the scalability and availability of the data, as EBS volumes are only accessible within a single availability zone and have a maximum size of 16 TiB. Also, EBS volumes are more expensive than S3 buckets and require provisioning and managing EC2 instances.

* Storing datasets as tables in a multi-node Amazon Redshift cluster would incur higher costs and complexity than using S3 and Athena. Amazon Redshift is a data warehouse service that is optimized for analytical queries over structured or semi-structured data. However, it requires setting up and maintaining a cluster of nodes, loading data into tables, and choosing the right distribution and sort keys for optimal performance. Moreover, Amazon Redshift charges for both storage and compute, while S3 and Athena only charge for the amount of data stored and scanned, respectively.

* Storing datasets as global tables in Amazon DynamoDB would not be feasible for large amounts of data, as DynamoDB is a key-value and document database service that is designed for fast and consistent performance at any scale. However, DynamoDB has a limit of 400 KB per item and 25 GB per partition key value, which may not be enough for storing large datasets. Also, DynamoDB does not support SQL queries natively, and would require using a service like Amazon EMR or AWS Glue to run SQL queries over DynamoDB data.

Amazon S3 - Cloud Object Storage

Amazon Athena - Interactive SQL Queries for Data in Amazon S3

Amazon EBS - Amazon Elastic Block Store (EBS)

Amazon Redshift - Data Warehouse Solution - AWS

Amazon DynamoDB - NoSQL Cloud Database Service

NEW QUESTION # 192

A media company with a very large archive of unlabeled images, text, audio, and video footage wishes to index its assets to allow rapid identification of relevant content by the Research team. The company wants to use machine learning to accelerate the efforts of its in-house researchers who have limited machine learning expertise.

Which is the FASTEST route to index the assets?

- A. Use the AWS Deep Learning AMI and Amazon EC2 GPU instances to create custom models for audio transcription and topic modeling, and use object detection to tag data into distinct categories/classes.
- **B. Use Amazon Rekognition, Amazon Comprehend, and Amazon Transcribe to tag data into distinct categories/classes.**
- C. Use Amazon Transcribe to convert speech to text. Use the Amazon SageMaker Neural Topic Model (NTM) and Object Detection algorithms to tag data into distinct categories/classes.
- D. Create a set of Amazon Mechanical Turk Human Intelligence Tasks to label all footage.

Answer: B

Explanation:

Explanation

Amazon Rekognition, Amazon Comprehend, and Amazon Transcribe are AWS machine learning services that can analyze and extract metadata from images, text, audio, and video content. These services are easy to use, scalable, and do not require any machine learning expertise. They can help the media company to quickly index its assets and enable rapid identification of relevant content by the research team. Using these services is the fastest route to index the assets, compared to the other options that involve human intervention, custom model development, or additional steps. References:

AWS Media Intelligence Solutions

AWS Machine Learning Services

The Best Services For Running Machine Learning Models On AWS

NEW QUESTION # 193

A Machine Learning Specialist working for an online fashion company wants to build a data ingestion solution for the company's Amazon S3-based data lake.

The Specialist wants to create a set of ingestion mechanisms that will enable future capabilities comprised of:

- * Real-time analytics
- * Interactive analytics of historical data
- * Clickstream analytics
- * Product recommendations

Which services should the Specialist use?

- **A. AWS Glue as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for real-time data insights; Amazon Kinesis Data Firehose for delivery to Amazon ES for clickstream analytics; Amazon EMR to generate personalized product recommendations**
- B. Amazon Athena as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for historical data insights; Amazon DynamoDB streams for clickstream analytics; AWS Glue to generate personalized product recommendations
- C. Amazon Athena as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for near-realtime data insights; Amazon Kinesis Data Firehose for clickstream analytics; AWS Glue to generate personalized product recommendations
- D. AWS Glue as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for historical data insights; Amazon Kinesis Data Firehose for delivery to Amazon ES for clickstream analytics; Amazon EMR to generate personalized product recommendations

Answer: A

Explanation:

The best services to use for building a data ingestion solution for the company's Amazon S3-based data lake are:

AWS Glue as the data catalog: AWS Glue is a fully managed extract, transform, and load (ETL) service that can discover, crawl, and catalog data from various sources and formats, and make it available for analysis. AWS Glue can also generate ETL code in Python or Scala to transform, enrich, and join data using AWS Glue Data Catalog as the metadata repository. AWS Glue Data Catalog is a central metadata store that integrates with Amazon Athena, Amazon EMR, and Amazon Redshift Spectrum, allowing users to create a unified view of their data across various sources and formats.

Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for real-time data insights: Amazon Kinesis Data Streams is a service that enables users to collect, process, and analyze real-time streaming data at any scale. Users can create data streams that can capture data from various sources, such as web and mobile applications, IoT devices, and social media platforms. Amazon Kinesis Data Analytics is a service that allows users to analyze streaming data using standard SQL queries or Apache Flink applications. Users can create real-time dashboards, metrics, and alerts based on the streaming data analysis results.

Amazon Kinesis Data Firehose for delivery to Amazon ES for clickstream analytics: Amazon Kinesis Data Firehose is a service that

enables users to load streaming data into data lakes, data stores, and analytics services. Users can configure Kinesis Data Firehose to automatically deliver data to various destinations, such as Amazon S3, Amazon Redshift, Amazon OpenSearch Service, and third-party solutions. For clickstream analytics, users can use Kinesis Data Firehose to deliver data to Amazon OpenSearch Service, a fully managed service that offers search and analytics capabilities for log data. Users can use Amazon OpenSearch Service to perform interactive analysis and visualization of clickstream data using Kibana, an open-source tool that is integrated with Amazon OpenSearch Service.

Amazon EMR to generate personalized product recommendations: Amazon EMR is a service that enables users to run distributed data processing frameworks, such as Apache Spark, Apache Hadoop, and Apache Hive, on scalable clusters of EC2 instances. Users can use Amazon EMR to perform advanced analytics, such as machine learning, on large and complex datasets stored in Amazon S3 or other sources. For product recommendations, users can use Amazon EMR to run Spark MLlib, a library that provides scalable machine learning algorithms, such as collaborative filtering, to generate personalized recommendations based on user behavior and preferences.

References:

AWS Glue - Fully Managed ETL Service

Amazon Kinesis - Data Streaming Service

Amazon OpenSearch Service - Managed OpenSearch Service

Amazon EMR - Managed Hadoop Framework

NEW QUESTION # 194

A Machine Learning Specialist is configuring automatic model tuning in Amazon SageMaker. When using the hyperparameter optimization feature, which of the following guidelines should be followed to improve optimization?

Choose the maximum number of hyperparameters supported by

- A. Execute only one hyperparameter tuning job at a time and improve tuning through successive rounds of experiments
- B. Specify a very large hyperparameter range to allow Amazon SageMaker to cover every possible value.
- C. Use log-scaled hyperparameters to allow the hyperparameter space to be searched as quickly as possible
- D. Amazon SageMaker to search the largest number of combinations possible

Answer: C

Explanation:

Explanation

Using log-scaled hyperparameters is a guideline that can improve the automatic model tuning in Amazon SageMaker. Log-scaled hyperparameters are hyperparameters that have values that span several orders of magnitude, such as learning rate, regularization parameter, or number of hidden units. Log-scaled hyperparameters can be specified by using a log-uniform distribution, which assigns equal probability to each order of magnitude within a range. For example, a log-uniform distribution between 0.001 and 1000 can sample values such as 0.001, 0.01, 0.1, 1, 10, 100, or 1000 with equal probability. Using log-scaled hyperparameters can allow the hyperparameter optimization feature to search the hyperparameter space more efficiently and effectively, as it can explore different scales of values and avoid sampling values that are too small or too large. Using log-scaled hyperparameters can also help avoid numerical issues, such as underflow or overflow, that may occur when using linear-scaled hyperparameters. Using log-scaled hyperparameters can be done by setting the `ScalingType` parameter to `Logarithmic` when defining the hyperparameter ranges in Amazon SageMaker^{1,2}. The other options are not valid or relevant guidelines for improving the automatic model tuning in Amazon SageMaker. Choosing the maximum number of hyperparameters supported by Amazon SageMaker to search the largest number of combinations possible is not a good practice, as it can increase the time and cost of the tuning job and make it harder to find the optimal values. Amazon SageMaker supports up to 20 hyperparameters for tuning, but it is recommended to choose only the most important and influential hyperparameters for the model and algorithm, and use default or fixed values for the rest³. Specifying a very large hyperparameter range to allow Amazon SageMaker to cover every possible value is not a good practice, as it can result in sampling values that are irrelevant or impractical for the model and algorithm, and waste the tuning budget. It is recommended to specify a reasonable and realistic hyperparameter range based on the prior knowledge and experience of the model and algorithm, and use the results of the tuning job to refine the range if needed⁴. Executing only one hyperparameter tuning job at a time and improving tuning through successive rounds of experiments is not a good practice, as it can limit the exploration and exploitation of the hyperparameter space and make the tuning process slower and less efficient. It is recommended to use parallelism and concurrency to run multiple training jobs simultaneously and leverage the Bayesian optimization algorithm that Amazon SageMaker uses to guide the search for the best hyperparameter values⁵.

NEW QUESTION # 195

.....

Reliable MLS-C01 Exam Registration: <https://www.updatedumps.com/Amazon/MLS-C01-updated-exam-dumps.html>

- 2026 Latest UpdateDumps MLS-C01 PDF Dumps and MLS-C01 Exam Engine Free Share: <https://drive.google.com/open?id=1fLdFvg8dlgBTYqIovsGaufokLy2A100>