# HOT AWS-Certified-Machine-Learning-Specialty Reliable Real Test - Latest Amazon AWS Certified Machine Learning - Specialty - Exam AWS-Certified-Machine-Learning-Specialty Dumps



BONUS!!! Download part of Real4exams AWS-Certified-Machine-Learning-Specialty dumps for free:
https://drive.google.com/open?id=1axXjbb_wshtTT8ooPEJJrgq-yIe-lZal

No doubt the Amazon AWS-Certified-Machine-Learning-Specialty certification exam is one of the most difficult Real4exams certification exams in the modern Real4exams world. This AWS-Certified-Machine-Learning-Specialty exam always gives a tough time to their candidates. The Real4exams understands this challenge and offers real, valid, and top-notch Amazon AWS-Certified-Machine-Learning-Specialty Exam Dumps in three different formats. All these three AWS-Certified-Machine-Learning-Specialty exam questions formats are easy to use and compatible with all devices, operating systems, and web browsers.

## Preparation Process

Many useful resources are available for the Amazon MLS-C01 exam. Let's take a closer look at them.

- **Discussion Forums**

  Besides the official training options offered by the vendor, many websites provide the candidates with the topic details and useful information related to the exam as well as the certification. Some of these sites include Quora and ExamLabs.com.

- **AWS Training Courses**

  The vendor offers the official Amazon AWS Machine Learning Certification Training, which you can access by visiting the Amazon website. It is good to know that these training courses are usually available as free.

- **AWS Machine Learning Documentation**

  The official AWS Machine Learning Documentation is one of the best preparation guides for the Amazon AWS Certified Machine Learning – Specialty exam. It is a great material to learn all the topics included in the exam content. You can find documentation, such as splitting data, data transformation, as well as types of data.

**>> AWS-Certified-Machine-Learning-Specialty Reliable Real Test <<**

## Free PDF Updated AWS-Certified-Machine-Learning-Specialty - AWS Certified Machine Learning - Specialty Reliable Real Test

Despite the complex technical concepts, our AWS-Certified-Machine-Learning-Specialty exam questions have been simplified to the level of average candidates, posing no hurdles in understanding the various ideas. It is also the reason that our AWS-Certified-Machine-Learning-Specialty study guide is famous all over the world. We also have tens of thousands of our loyal customers who

support us on the AWS-Certified-Machine-Learning-Specialty Learning Materials. Just look at the feedbacks on our website, they all praised our AWS-Certified-Machine-Learning-Specialty practice engine.

# AWS Machine Learning Specialty Exam Syllabus Topics:

| Section | Objectives |
|---|---|
| **Data Engineering - 20%** | |
| Create data repositories for machine learning. | - Identify data sources (e.g., content and location, primary sources such as user data)<br>- Determine storage mediums (e.g., DB, Data Lake, S3, EFS, EBS) |
| Identify and implement a data ingestion solution. | - Data job styles/types (batch load, streaming)<br><br>• Kinesis<br>• Kinesis Analytics<br>• Kinesis Firehose<br>• EMR<br>• Glue<br><br>- Data ingestion pipelines (Batch-based ML workloads and streaming-based ML workloads)<br>- Job scheduling |
| Identify and implement a data transformation solution. | - Transforming data transit (ETL: Glue, EMR, AWS Batch)<br>- Handle ML-specific data using map reduce (Hadoop, Spark, Hive) |
| **Exploratory Data Analysis - 24%** | |
| Sanitize and prepare data for modeling. | - Identify and handle missing data, corrupt data, stop words, etc.<br>- Formatting, normalizing, augmenting, and scaling data<br>- Labeled data (recognizing when you have enough labeled data and identifying mitigation strategies [Data labeling tools (Mechanical Turk, manual labor)]) |
| Perform feature engineering. | - Identify and extract features from data sets, including from data sources such as text, speech, image, public datasets, etc.<br>- Analyze/evaluate feature engineering concepts (binning, tokenization, outliers, synthetic features, 1 hot encoding, reducing dimensionality of data) |
| Analyze and visualize data for machine learning. | - Graphing (scatter plot, time series, histogram, box plot)<br>- Interpreting descriptive statistics (correlation, summary statistics, p value)<br>- Clustering (hierarchical, diagnosing, elbow plot, cluster size) |
| **Modeling - 36%** | |
| Frame business problems as machine learning problems. | - Determine when to use/when not to use ML<br>- Know the difference between supervised and unsupervised learning<br>- Selecting from among classification, regression, forecasting, clustering, recommendation, etc. |
| Select the appropriate model(s) for a given machine learning problem. | - Xgboost, logistic regression, K-means, linear regression, decision trees, random forests, RNN, CNN, Ensemble, Transfer learning<br>- Express intuition behind models |
| Train machine learning models. | - Train validation test split, cross-validation<br>- Optimizer, gradient descent, loss functions, local minima, convergence, batches, probability, etc.<br>- Compute choice (GPU vs. CPU, distributed vs. non-distributed, platform [Spark vs. non-Spark])<br>- Model updates and retraining<br><br>• Batch vs. real-time/online |

| | |
|---|---|
| Perform hyperparameter optimization. | - Regularization<br><br> • Drop out<br> • L1/L2<br><br>- Cross validation<br>- Model initialization<br>- Neural network architecture (layers/nodes), learning rate, activation functions<br>- Tree-based models (# of trees, # of levels)<br>- Linear models (learning rate) |

# Amazon AWS Certified Machine Learning - Specialty Sample Questions (Q231-Q236):

NEW QUESTION # 231
A real-estate company is launching a new product that predicts the prices of new houses. The historical data for the properties and prices is stored in .csv format in an Amazon S3 bucket. The data has a header, some categorical fields, and some missing values. The company's data scientists have used Python with a common open-source library to fill the missing values with zeros. The data scientists have dropped all of the categorical fields and have trained a model by using the open-source linear regression algorithm with the default parameters.
The accuracy of the predictions with the current model is below 50%. The company wants to improve the model performance and launch the new product as soon as possible.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Create an IAM role with access to Amazon S3, Amazon SageMaker, and AWS Lambda. Create a training job with the SageMaker built-in XGBoost model pointing to the bucket with the dataset.
  Specify the price as the target feature. Wait for the job to complete. Load the model artifact to a Lambda function for inference on prices of new houses.
- B. Create an IAM role for Amazon SageMaker with access to the S3 bucket. Create a SageMaker AutoML job with SageMaker Autopilot pointing to the bucket with the dataset. Specify the price as the target attribute. Wait for the job to complete. Deploy the best model for predictions.
- C. Create an Amazon SageMaker notebook with a new IAM role that is associated with the notebook. Pull the dataset from the S3 bucket. Explore different combinations of feature engineering transformations, regression algorithms, and hyperparameters. Compare all the results in the notebook, and deploy the most accurate configuration in an endpoint for predictions.
- D. Create a service-linked role for Amazon Elastic Container Service (Amazon ECS) with access to the S3 bucket. Create an ECS cluster that is based on an AWS Deep Learning Containers image. Write the code to perform the feature engineering. Train a logistic regression model for predicting the price, pointing to the bucket with the dataset. Wait for the training job to complete. Perform the inferences.

Answer: B

Explanation:
The solution D meets the requirements with the least operational overhead because it uses Amazon SageMaker Autopilot, which is a fully managed service that automates the end-to-end process of building, training, and deploying machine learning models. Amazon SageMaker Autopilot can handle data preprocessing, feature engineering, algorithm selection, hyperparameter tuning, and model deployment. The company only needs to create an IAM role for Amazon SageMaker with access to the S3 bucket, create a SageMaker AutoML job pointing to the bucket with the dataset, specify the price as the target attribute, and wait for the job to complete. Amazon SageMaker Autopilot will generate a list of candidate models with different configurations and performance metrics, and the company can deploy the best model for predictions1.
The other options are not suitable because:
* Option A: Creating a service-linked role for Amazon Elastic Container Service (Amazon ECS) with access to the S3 bucket, creating an ECS cluster based on an AWS Deep Learning Containers image, writing the code to perform the feature engineering, training a logistic regression model for predicting the price, and performing the inferences will incur more operational overhead than using Amazon SageMaker Autopilot. The company will have to manage the ECS cluster, the container image, the code, the model, and the inference endpoint. Moreover, logistic regression may not be the best algorithm for predicting the price, as it is more suitable for binary classification tasks2.
* Option B: Creating an Amazon SageMaker notebook with a new IAM role that is associated with the notebook, pulling the dataset from the S3 bucket, exploring different combinations of feature engineering transformations, regression algorithms, and

hyperparameters, comparing all the results in the notebook, and deploying the most accurate configuration in an endpoint for predictions will incur more operational overhead than using Amazon SageMaker Autopilot. The company will have to write the code for the feature engineering, the model training, the model evaluation, and the model deployment. The company will also have to manually compare the results and select the best configuration3.
* Option C: Creating an IAM role with access to Amazon S3, Amazon SageMaker, and AWS Lambda, creating a training job with the SageMaker built-in XGBoost model pointing to the bucket with the dataset, specifying the price as the target feature, loading the model artifact to a Lambda function for inference on prices of new houses will incur more operational overhead than using Amazon SageMaker Autopilot. The company will have to create and manage the Lambda function, the model artifact, and the inference endpoint. Moreover, XGBoost may not be the best algorithm for predicting the price, as it is more suitable for classification and ranking tasks4.
1: Amazon SageMaker Autopilot
2: Amazon Elastic Container Service
3: Amazon SageMaker Notebook Instances
4: Amazon SageMaker XGBoost Algorithm

# NEW QUESTION # 232
A company is using Amazon SageMaker to build a machine learning (ML) model to predict customer churn based on customer call transcripts. Audio files from customer calls are located in an on-premises VoIP system that has petabytes of recorded calls. The on-premises infrastructure has high-velocity networking and connects to the company's AWS infrastructure through a VPN connection over a 100 Mbps connection.
The company has an algorithm for transcribing customer calls that requires GPUs for inference. The company wants to store these transcriptions in an Amazon S3 bucket in the AWS Cloud for model development.
Which solution should an ML specialist use to deliver the transcriptions to the S3 bucket as quickly as possible?

* A. Order and use AWS Outposts to run the transcription algorithm on GPU-based Amazon EC2 instances. Store the resulting transcriptions in the transcription S3 bucket.
* B. Order and use an AWS Snowball Edge Compute Optimized device with an NVIDIA Tesla module to run the transcription algorithm. Use AWS DataSync to send the resulting transcriptions to the transcription S3 bucket.
* C. Order and use an AWS Snowcone device with Amazon EC2 Inf1 instances to run the transcription algorithm Use AWS DataSync to send the resulting transcriptions to the transcription S3 bucket
* D. Use AWS DataSync to ingest the audio files to Amazon S3. Create an AWS Lambda function to run the transcription algorithm on the audio files when they are uploaded to Amazon S3. Configure the function to write the resulting transcriptions to the transcription S3 bucket.

**Answer: B**

Explanation:
The company needs to transcribe petabytes of audio files from an on-premises VoIP system to an S3 bucket in the AWS Cloud. The transcription algorithm requires GPUs for inference, which are not available on the on- premises system. The VPN connection over a 100 Mbps connection is not sufficient to transfer the large amount of data quickly. Therefore, the company should use an AWS Snowball Edge Compute Optimized device with an NVIDIA Tesla module to run the transcription algorithm locally and leverage the GPU power.
The device can store up to 42 TB of data and can be shipped back to AWS for data ingestion. The company can use AWS DataSync to send the resulting transcriptions to the transcription S3 bucket in the AWS Cloud.
This solution minimizes the network bandwidth and latency issues and enables faster data processing and transfer.
Option B is incorrect because AWS Snowcone is a small, portable, rugged, and secure edge computing and data transfer device that can store up to 8 TB of data. It is not suitable for processing petabytes of data and does not support GPU-based instances.
Option C is incorrect because AWS Outposts is a service that extends AWS infrastructure, services, APIs, and tools to virtually any data center, co-location space, or on-premises facility. It is not designed for data transfer and ingestion, and it would require additional infrastructure and maintenance costs.
Option D is incorrect because AWS DataSync is a service that makes it easy to move large amounts of data to and from AWS over the internet or AWS Direct Connect. However, using DataSync to ingest the audio files to S3 would still be limited by the network bandwidth and latency. Moreover, running the transcription algorithm on AWS Lambda would incur additional costs and complexity, and it would not leverage the GPU power that the algorithm requires.
References:
* AWS Snowball Edge Compute Optimized
* AWS DataSync
* AWS Snowcone
* AWS Outposts
* AWS Lambda

**NEW QUESTION # 233**

A Machine Learning Specialist is implementing a full Bayesian network on a dataset that describes public transit in New York City.
One of the random variables is discrete, and represents the number of minutes New Yorkers wait for a bus given that the buses
cycle every 10 minutes, with a mean of 3 minutes.
Which prior probability distribution should the ML Specialist use for this variable?

- A. Poisson distribution ,
- B. Uniform distribution
- C. Binomial distribution
- D. Normal distribution

**Answer: C**


**NEW QUESTION # 234**

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. Amazon Athena as the data catalog: Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for near-real-time
  data insights; Amazon Kinesis Data Firehose for clickstream analytics; AWS Glue 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. 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
- D. 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

**Answer: D**


**NEW QUESTION # 235**

A Data Science team within a large company uses Amazon SageMaker notebooks to access data stored in Amazon S3 buckets.
The IT Security team is concerned that internet-enabled notebook instances create a security vulnerability where malicious code
running on the instances could compromise data privacy. The company mandates that all instances stay within a secured VPC with
no internet access, and data communication traffic must stay within the AWS network.
How should the Data Science team configure the notebook instance placement to meet these requirements?

- A. Associate the Amazon SageMaker notebook with a private subnet in a VPC. Ensure the VPC has S3 VPC endpoints and
  Amazon SageMaker VPC endpoints attached to it.
- B. Associate the Amazon SageMaker notebook with a private subnet in a VPC. Place the Amazon SageMaker endpoint and
  S3 buckets within the same VPC.
- C. Associate the Amazon SageMaker notebook with a private subnet in a VPC. Ensure the VPC has a NAT gateway and an
  associated security group allowing only outbound connections to Amazon S3 and Amazon SageMaker
- D. Associate the Amazon SageMaker notebook with a private subnet in a VPC. Use 1AM policies to grant access to
  Amazon S3 and Amazon SageMaker.

**Answer: A**

Explanation:

Explanation

To configure the notebook instance placement to meet the requirements, the Data Science team should associate the Amazon SageMaker notebook with a private subnet in a VPC. A VPC is a virtual network that is logically isolated from other networks in AWS. A private subnet is a subnet that has no internet gateway attached to it, and therefore cannot communicate with the internet. By placing the notebook instance in a private subnet, the team can ensure that it stays within a secured VPC with no internet access. However, to access data stored in Amazon S3 buckets and other AWS services, the team needs to ensure that the VPC has S3 VPC endpoints and Amazon SageMaker VPC endpoints attached to it. A VPC endpoint is a gateway that enables private connections between the VPC and supported AWS services. A VPC endpoint does not require an internet gateway, a NAT device, or a VPN connection, and ensures that the traffic between the VPC and the AWS service does not leave the AWS network. By using VPC endpoints, the team can access Amazon S3 and Amazon SageMaker from the notebook instance without compromising data privacy or security.

References:

1: What Is Amazon VPC? - Amazon Virtual Private Cloud

2: Subnet Routing - Amazon Virtual Private Cloud

3: VPC Endpoints - Amazon Virtual Private Cloud


**NEW QUESTION # 236**

......

**Exam AWS-Certified-Machine-Learning-Specialty Dumps**: https://www.real4exams.com/AWS-Certified-Machine-Learning-Specialty_braindumps.html

Braindumps □ AWS-Certified-Machine-Learning-Specialty Reliable Test Objectives □ Search for ⇒ AWS-Certified-Machine-Learning-Specialty ⇐ and obtain a free download on [ www.examcollectionpass.com ] □AWS-Certified-Machine-Learning-Specialty Latest Test Question

- www.stes.tyc.edu.tw, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, ntc-israel.com, wmwimal.com, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, Disposable vapes

BTW, DOWNLOAD part of Real4exams AWS-Certified-Machine-Learning-Specialty dumps from Cloud Storage: https://drive.google.com/open?id=1axXjbb_wshtTT8ooPEJJrgq-yIe-lZal