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.leads4pass.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

An e commerce company wants to launch a new cloud-based product recommendation feature for its web application. Due to data localization regulations, any sensitive data must not leave its on-premises data center, and the product recommendation model must be trained and tested using nonsensitive data only. Data transfer to the cloud must use IPsec. The web application is hosted on premises with a PostgreSQL database that contains all the data. The company wants the data to be uploaded securely to Amazon S3 each day for model retraining.

How should a machine learning specialist meet these requirements?

A. Create an AWS Glue job to connect to the PostgreSQL DB instance. Ingest tables without sensitive data through an AWS Site-to-Site VPN connection directly into Amazon S3.

B. Create an AWS Glue job to connect to the PostgreSQL DB instance. Ingest all data through an AWS Site-to-Site VPN connection into Amazon S3 while removing sensitive data using a PySpark job.

C. Use AWS Database Migration Service (AWS DMS) with table mapping to select PostgreSQL tables with no sensitive data through an SSL connection. Replicate data directly into Amazon S3.

D. Use PostgreSQL logical replication to replicate all data to PostgreSQL in Amazon EC2 through AWS Direct Connect with a VPN connection. Use AWS Glue to move data from Amazon EC2 to Amazon S3.

Correct Answer: A

This solution meets the requirements of data localization regulations and secure data transfer. By creating an AWS Glue job to connect to the PostgreSQL DB instance, the machine learning specialist can extract tables without sensitive data. By using a Site-to-Site VPN connection, the data can be securely transferred from the on-premises data center to Amazon S3, where it can be used for model retraining. This solution ensures that any sensitive data remains in the on-premises data center, and that only non-sensitive data is uploaded to the cloud.

https://docs.aws.amazon.com/glue/latest/dg/aws-glue-programming-etl-connect.html#aws-glue-programming-etl-connect-jdbc https://aws.amazon.com/blogs/big-data/how-to-access-and-analyze-on-premises-data-stores-using-aws-glue/

QUESTION 2

A geospatial analysis company processes thousands of new satellite images each day to produce vessel detection data for commercial shipping. The company stores the training data in Amazon S3. The training data incrementally increases in size with new images each day.

The company has configured an Amazon SageMaker training job to use a single ml.p2.xlarge instance with File input mode to train the built-in Object Detection algorithm. The training process was successful last month but is now failing because of a lack of storage. Aside from the addition of training data, nothing has changed in the model training process.

A machine learning (ML) specialist needs to change the training configuration to fix the problem. The solution must optimize performance and must minimize the cost of training.

Which solution will meet these requirements?

A. Modify the training configuration to use two ml.p2.xlarge instances.

B. Modify the training configuration to use Pipe input mode.

C. Modify the training configuration to use a single ml.p3.2xlarge instance.

D. Modify the training configuration to use Amazon Elastic File System (Amazon EFS) instead of Amazon S3 to store the input training data.

Correct Answer: B

https://aws.amazon.com/blogs/machine-learning/using-pipe-input-mode-for-amazon-sagemaker-algorithms/

QUESTION 3

A retail company intends to use machine learning to categorize new products A labeled dataset of current products was provided to the Data Science team The dataset includes 1 200 products The labeled dataset has 15 features for each product such as title dimensions, weight, and price Each product is labeled as belonging to one of six categories such as books, games, electronics, and movies.

Which model should be used for categorizing new products using the provided dataset for training?

A. An XGBoost model where the objective parameter is set to multi: softmax

B. A deep convolutional neural network (CNN) with a softmax activation function for the last layer

C. A regression forest where the number of trees is set equal to the number of product categories

D. A DeepAR forecasting model based on a recurrent neural network (RNN)

Correct Answer: A

The XGBoost algorithm is a popular and effective technique for multi-class classification. The objective parameter can be set to multi:softmax, which uses a softmax objective function for multi-class classification. This will train the model to

predict the probability of each product belonging to each category, and the most probable category will be chosen as the final prediction.

A deep convolutional neural network (CNN) (B) is a powerful technique commonly used for image recognition tasks. However, it is less appropriate for tabular data like the dataset provided.

https://medium.com/@gabrielziegler3/multiclass-multilabel-classification-with-xgboost-66195e4d9f2d

QUESTION 4

This graph shows the training and validation loss against the epochs for a neural network
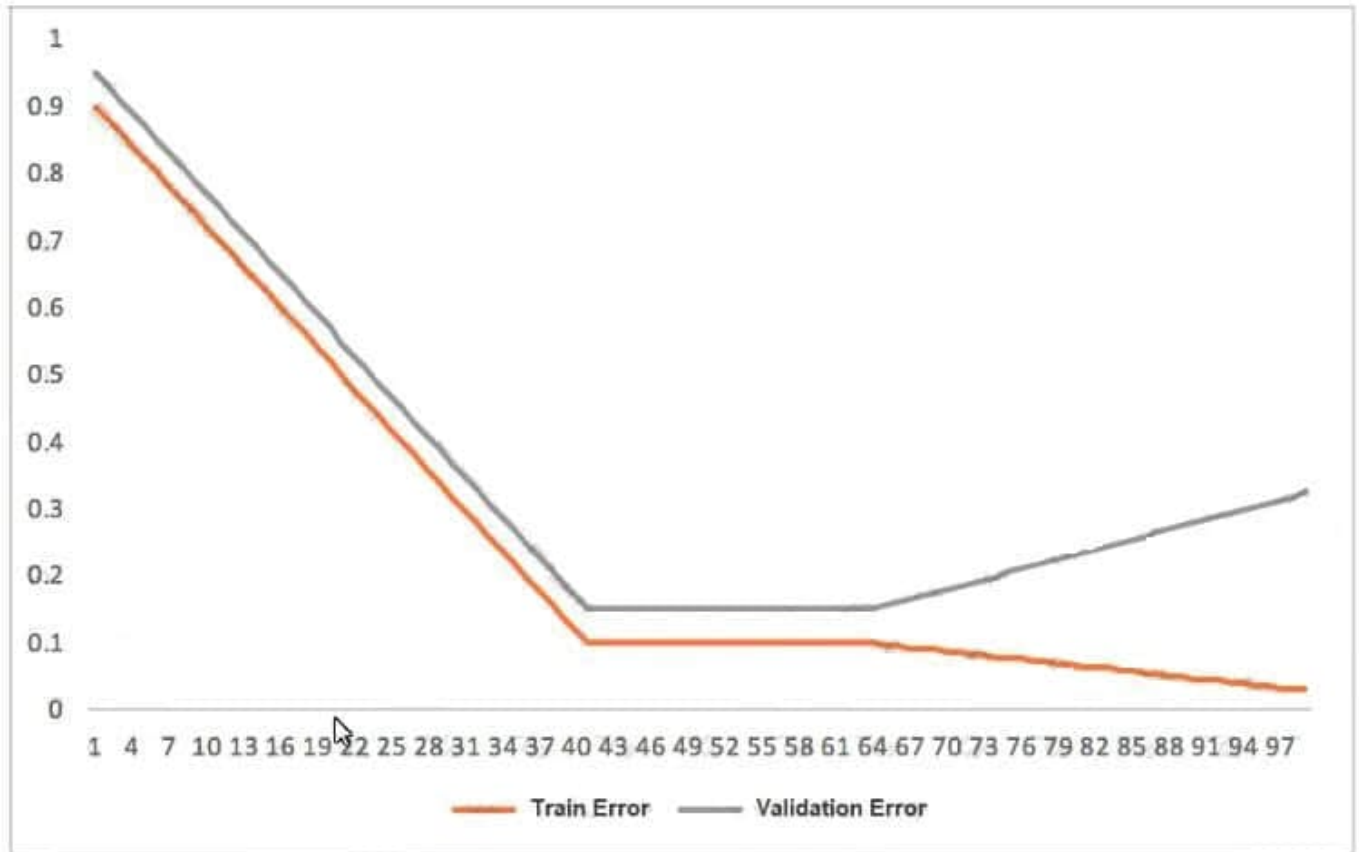
The network being trained is as follows

1.

 Two dense layers one output neuron

2.

 100 neurons in each layer

3.

100 epochs

4.

Random initialization of weights

Which technique can be used to improve model performance in terms of accuracy in the validation set?

A. Early stopping

B. Random initialization of weights with appropriate seed

C. Increasing the number of epochs

D. Adding another layer with the 100 neurons

Correct Answer: A

The answer is Early Stopping. Stopp the training before accuracy start do decrease.

QUESTION 5

A large JSON dataset for a project has been uploaded to a private Amazon S3 bucket The Machine Learning Specialist wants to securely access and explore the data from an Amazon SageMaker notebook instance A new VPC was created and assigned to the Specialist

How can the privacy and integrity of the data stored in Amazon S3 be maintained while granting access to the Specialist for analysis?

A. Launch the SageMaker notebook instance within the VPC with SageMaker-provided internet access enabled Use an S3 ACL to open read privileges to the everyone group

B. Launch the SageMaker notebook instance within the VPC and create an S3 VPC endpoint for the notebook to access the data Copy the JSON dataset from Amazon S3 into the ML storage volume on the SageMaker notebook instance and work against the local dataset

C. Launch the SageMaker notebook instance within the VPC and create an S3 VPC endpoint for the notebook to access the data Define a custom S3 bucket policy to only allow requests from your VPC to access the S3 bucket

D. Launch the SageMaker notebook instance within the VPC with SageMaker-provided internet access enabled. Generate an S3 pre-signed URL for access to data in the bucket

Correct Answer: B

[MLS-C01 VCE Dumps](#)   [MLS-C01 Study Guide](#)   [MLS-C01 Exam Questions](#)