

AWS Machine Learning Specialty Practice Exam - Quiz

Q1) .

- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .
- .

Answer Sheet

Q2) A legacy MapReduce job is running on EMR, and must ingest data that has been moved from HDFS to S3. Which is a viable option for connecting the S3 data to MapReduce on EMR?

Use EFS to connect MapReduce to the S3 bucket.

MapReduce can talk natively to S3 using the s3a:// prefix

Use EMRFS to connect MapReduce to S3, using the s3:// file prefix.

Explanation: -EMR extends Hadoop (which includes MapReduce) to use S3 as a storage backend instead of HDFS. Use Apache Hive as an intermediary between MapReduce and S3

Q3) Your deep neural network has achieved 99% accuracy on your training data, but only 90% on your test data. The model's performance for the same task is around 98%.

What are some possible conclusions (SELECT TWO)?

The model is overfitting and failing to generalize. Using more layers may help.

The model is overfitting and failing to generalize. Using dropout regularization may help.

Explanation: -You'll need to understand how to deal with overfitting through regularization. Dropout regularization is one way to prevent overfitting in a neural network. Shuffling the input data every epoch would also help.

The model is underfitting, and broader hidden layers are needed.

The model is overfitting and failing to generalize. Using fewer layers may help.

Q4) A media company wishes to recommend movies to users based on the predicted rating of that movie. They are using SageMaker and Factorization Machines.

What format should the rating data used for training the model be in?

CSV

RecordIO/ protobuf in float32 format