QUESTION 211

A company is deploying a new version of a classification model for an online service using Amazon Sagemaker .Service utilization is close to 24/7 , and there are no maintenance windows defined to deploy and test the new model .Company executives are concerned about the potential impact on user experience if the new model does not behave as expected in production.

What is the SIMPLEST way to deploy the new model on Amazon Sagemaker to reduce the potential impact on user experience ?

A. Deploy the new model as a production var ant on the same endpoint as the existing model conure an initial traffic flow of 50% , and then gradually decrease the traffic flow

B. Deploy the new model on an endpoint and distribute user requests with an Elastic Load Balancer .

C. Deploy the new model as a production variant on the same endpoint as the existing model , conure an initial traffic low of 10 , and then gradually increase the trafic flow

D. Deploy the new model on an endpoint and distribute user requests with Amazon Cloudfront

Answer:D

QUESTION 212

For the given confusion matrix , what is the recall and precision of the model ?

		Actual	
		Yes	No
	Yes	12	3
Predicted	No	1	9

A. Recall = 0.92 Precision 0.84
B. Recall = 0.84 Precision 0.8
C. Recall = 0.92 Precision 0.8
D. Recall = 0.8 Precision 0.92
Answer:C

QUESTION 213

A machine learning specialist is raining a convolutional neural network (CNN) on a cluster of Amazon EC2 P3 instances for a computer vision task .However , training is taking longer than expected .The specialist tried reducing the resolution of images to speed up training , but the accuracy dropped below an acceptable level.

Which technique may speed up training without significantly degrading accuracy ?(Select Two)

A. Use mixed precision.
B. Replace convolutional layers with fully connected layers
C. Start with a pretrained CNN and freeze most layers during training
D. Replace the P3 instances with G3 instances
E. Downsample training data for each class while keeping the class ratios balanced .
Answer:AC

QUESTION 214
A data scientist is building a logistic regression model for a classification , but has not been able to achieve acceptable accuracy .The scientist feels that the feature space is not linearly separable .There is a strong business need to build an interpretable model where the contribution of each feature in the classification must be well understood .The inference must begenerated as quickly as possible.Which measures should the data scientist take to achieve these goals with the LEAST amount of effort ?
A. Apply nonlinear transformations to the existing features
B. Remove features from the model one at a time using recursive feature elimination
C. Create a multilayer perceptron(MLP) with one or more hidden layers
D. Create a linear regression model
Answer:B

QUESTION 215
A data scientist is analzing a manufacuing dataset to datermine wherther a product assemled a various facory locations will pass or fail quality inspections.Each obsevation includes a factory ID , various aspects of the profuct ,and whether it passed or failed the quality inspection.Thirty percent of the observations are missing the factory ID , which is problematic because different factories appear to have diferent failue rates .The data scientist wants to predict whether the product will pass inspection to automate the inspection process.
Which approach can manage observations that are missing the factory ID ?
A. Drop the dataset rows corresponding to the missing factory IDs .
B. Impute the factory ID as the mode of the factory ID values
C. Impute the factory ID as the mean of the factory ID values
D. Model the missing factory ID as a value of a categorical or substitute attribute
Answer:C

QUESTION 216
A company has asked a machine learning specialist to increase the efficiency of its direct advertisement campaigns .Curently , there are no mechanisms in place to group customers , so campaigns are sent to the whole customer base .This is becoming too expensive as the customer base continues to grow.
The company has provided the specialist with access to a database containing 5 year's worth of purchasing data.
what approach should the machine learning specialist take to achieve the MOST precise results to minimize the number of mistargeted advertisements ?
A. Apply the k-means algorithm to create groups of similar customers based on their purchasing experience
B. Apply the factorizaton machine al hm o classify customers in different groups based on their purchasing experience
C. Create a Lambda function to query the database for the most common products .Use this output to feed an Xgboost algorithm to classify customers into different classes

D. Use an EMR pipeline to consolidate information for all past ansons and create oups of users by gene and ae .Use these as destination groups for the advertisements

Answer:B

QUESTION 217

A manufacturing company asks its machine learning specialist to develop a model that classifies defective parts into one of eight defect types.The company has provided roughly 100,000 images per defect type for training.During the initial training of the image classification model ,the specialist notices that the validation accuracy is 80%,while the training accuracy is 90%,it is known that human-level performance for this type of image classification is around 90%.

What should the specialist consider to fix this issue?

A. A longer training time
B. Making the network larger
C. Using a different optimizer
D. Using some form of regularization

Answer:A

QUESTION 218

A company is developing a custom virtual voice assistant .The first version of the virtual voice assistant simply asks the user for their name and then asks a simple question .The response will be analyzed using sentiment analysis before the voice assistant responds to the user utilizing speech .Amazon Polly will be used to output the response as speech.The company wants to use an entirely serverless architecture levering managed AWS services with the goal of having a minimal infrastructure configuration.

Which AWS services or features should the company leverage to build the system ?

A. Amazon Lex to manage the conversational interface , AWS Lambda to manage the serverless compute , and Amazon Comprehend to analyze if the comment is positive or negative

B. Amazon Transcribe to convert the speech into text , Amazon Connect to manage the conversational interface , Amazon ECS to manage the serverless compute , and Amazon Comprehend to analyze if the comment is positive or negative

C. Amazon Lex to manage the conversational interface , AWS Lambda to manage the serverless compute , and the Python Natural Language Toolkit running on AWS Lambda for sentiment analysis

D. Amazon Transcribe to convert the speech into text Amazon Lex to manage the conversational interface , AWS Lambda to manage the serverless compute , and AWS Comprehend to analyze if the comment is positive or negative

Answer:D

QUESTION 219

A machine leaning specialist needs to Sinstall packages on Amazon SageMaker-hosed notebooks that will be used to develop machine learning solutions .The specialist has several

custom packages and dependencies that need to remain constant and persistent across notebook environments.

Which of the following accomplishes this with the LEAST administrative overhead ?

A. Create a virtual environment using conda create from the Jupyter terminal on each notebook instance

B. Create a lifecycle configuration that triggers on notebook start to install the packages

C. Install the packages on the notebooks using the Jupyter terminal

D. Create a script that triggers the install of the packages during startup

Answer:B

QUESTION 220

A machine learning specialist has been using the Amazon Sagemaker automatic model tuning feature.In each tuning job , 4 hyperparameters are being tuned with total training jobs st to 30 and the parallel job training count set to 10 .Unfortunately , the objective metric achieved by the best training job identified by Amazon Sagemaker during tuning is insufficient for the use case.Which action is MOST likely to produce a tuning job with an improved objective metric for the best training job identified by Amazon SageMaker ?

A. Reduce the number of hyperparameters being tuned

B. Reduce the parallel job training count .

C. Change the objective metric .

D. Change the instance type used in the training jobs

Answer:A

QUESTION 221

A model trained with Amazon Sagemaker on a highly sensitive 10 TB dataset is delivering poor prediction results , even after several training jobs .Upon further review of the dataset .a machine learning specialist notices that some fields were pulled from the same data source and contain overlapping information.

How should the specialist clean up the data and securely access it for future training jobs ?

A. Use AWS Glue to discard the related fields and store the new dataset on Amazon S3 .Add an S3 VPC endpoint for Amazon Sagemaker for data access

B. Use Amazon EMR to discard the elated elds and store the new dataset on an Apache Hadoop Distributed File ystem(HDFS) .Add a private subnet in an EMR VPC for AmazonSagemaker for data access .

C. Use AWS Lambda to join the related fields and store the new dataset on Amazon Sagemaker Use Sagemaker EBS volumes for data access

D. Use Amazon Athena to join the related fields and store the new dataset on Amazon S3 .Add a NAT gateway for Amazon Sagemaker for data access

Answer:D