# [November-2020Braindump2go MLS-C01 PDF Dumps and MLS-C01 VCE Dumps Free Download[Q82-Q102

November/2020 Latest Braindump2go MLS-C01 Exam Dumps with PDF and VCE Free Updated Today! Following are some new MLS-C01 Real Exam Questions!QUESTION 82A Data Scientist is building a model to predict customer churn using a dataset of 100 continuous numerical features. The Marketing team has not provided any insight about which features are relevant for churn prediction. The Marketing team wants to interpret the model and see the direct impact of relevant features on the model outcome. While training a logistic regression model, the Data Scientist observes that there is a wide gap between the training and validation set accuracy.Which methods can the Data Scientist use to improve the model performance and satisfy the Marketing team's needs? (Choose two.)A.    Add L1 regularization to the classifierB.    Add features to the datasetC.    Perform recursive feature elimination D.    Perform t-distributed stochastic neighbor embedding (t-SNE)E.    Perform linear discriminant analysisAnswer: BEQUESTION 83An aircraft engine manufacturing company is measuring 200 performance metrics in a time-series. Engineers want to detect critical manufacturing defects in near-real time during testing. All of the data needs to be stored for offline analysis.What approach would be the MOST effective to perform near-real time defect detection?A.    Use AWS IoT Analytics for ingestion, storage, and further analysis. Use Jupyter notebooks from within AWS IoT Analytics to carry out analysis for anomalies.B.    Use Amazon S3 for ingestion, storage, and further analysis. Use an Amazon EMR cluster to carry out Apache Spark ML k-means clustering to determine anomalies.C.    Use Amazon S3 for ingestion, storage, and further analysis. Use the Amazon SageMaker Random Cut Forest (RCF) algorithm to determine anomalies.D.    Use Amazon Kinesis Data Firehose for ingestion and Amazon Kinesis Data Analytics Random Cut Forest (RCF) to perform anomaly detection. Use Kinesis Data Firehose to store data in Amazon S3 for further analysis.Answer: BQUESTION 84A Machine Learning team runs its own training algorithm on Amazon SageMaker. The training algorithm requires external assets. The team needs to submit both its own algorithm code and algorithm-specific parameters to Amazon SageMaker.What combination of services should the team use to build a custom algorithm in Amazon SageMaker? (Choose two.)A.    AWS Secrets ManagerB.    AWS CodeStarC.    Amazon ECRD.    Amazon ECSE.    Amazon S3Answer: CEQUESTION 85A Machine Learning Specialist wants to determine the appropriate SageMakerVariantInvocationsPerInstance setting for an endpoint automatic scaling configuration. The Specialist has performed a load test on a single instance and determined that peak requests per second (RPS) without service degradation is about 20 RPS. As this is the first deployment, the Specialist intends to set the invocation safety factor to 0.5.Based on the stated parameters and given that the invocations per instance setting is measured on a per-minute basis, what should the Specialist set as the SageMakerVariantInvocationsPerInstance setting?A.    10B.    30C.    600D.    2,400Answer: CQUESTION 86A company uses a long short-term memory (LSTM) model to evaluate the risk factors of a particular energy sector. The model reviews multi-page text documents to analyze each sentence of the text and categorize it as either a potential risk or no risk. The model is not performing well, even though the Data Scientist has experimented with many different network structures and tuned the corresponding hyperparameters.Which approach will provide the MAXIMUM performance boost?A.    Initialize the words by term frequency-inverse document frequency (TF-IDF) vectors pretrained on a large collection of news articles related to the energy sector.B.    Use gated recurrent units (GRUs) instead of LSTM and run the training process until the validation loss stops decreasing.C.    Reduce the learning rate and run the training process until the training loss stops decreasing.D.    Initialize the words by word2vec embeddings pretrained on a large collection of news articles related to the energy sector.Answer: CQUESTION 87A Machine Learning Specialist needs to move and transform data in preparation for training. Some of the data needs to be processed in near-real time, and other data can be moved hourly. There are existing Amazon EMR MapReduce jobs to clean and feature engineering to perform on the data.Which of the following services can feed data to the MapReduce jobs? (Choose two.)A.    AWS DMSB.    Amazon KinesisC.    AWS Data PipelineD.    Amazon AthenaE.    Amazon ESAnswer: AEQUESTION 88A Machine Learning Specialist previously trained a logistic regression model using scikit-learn on a local machine, and the Specialist now wants to deploy it to production for inference only.What steps should be taken to ensure Amazon SageMaker can host a model that was trained locally?A.    Build the Docker image with the inference code. Tag the Docker image with the registry hostname and upload it to Amazon ECR.B.    Serialize the trained model so the format is compressed for deployment. Tag the Docker image with the registry hostname and upload it to Amazon S3.C.    Serialize the trained model so the format is compressed for deployment. Build the image and upload it to Docker Hub.D.    Build the Docker image with the inference code. Configure Docker Hub and upload the image to Amazon ECR.Answer: DQUESTION 89A trucking company is collecting live image data from its fleet of trucks across the globe. The data is growing rapidly and approximately 100 GB of new data is generated every day. The company wants to explore machine learning uses cases while ensuring the data is only accessible to specific IAM users.Which storage option provides the most processing flexibility and will allow access control with IAM?A.    Use a database,

such as Amazon DynamoDB, to store the images, and set the IAM policies to restrict access to only the desired IAM users.B.    Use an Amazon S3-backed data lake to store the raw images, and set up the permissions using bucket policies.C.    Setup up Amazon EMR with Hadoop Distributed File System (HDFS) to store the files, and restrict access to the EMR instances using IAM policies. D.    Configure Amazon EFS with IAM policies to make the data available to Amazon EC2 instances owned by the IAM users.Answer: CQUESTION 90A credit card company wants to build a credit scoring model to help predict whether a new credit card applicant will default on a credit card payment. The company has collected data from a large number of sources with thousands of raw attributes. Early experiments to train a classification model revealed that many attributes are highly correlated, the large number of features slows down the training speed significantly, and that there are some overfitting issues.The Data Scientist on this project would like to speed up the model training time without losing a lot of information from the original dataset.Which feature engineering technique should the Data Scientist use to meet the objectives?A.    Run self-correlation on all features and remove highly correlated featuresB.    Normalize all numerical values to be between 0 and 1C.    Use an autoencoder or principal component analysis (PCA) to replace original features with new featuresD.    Cluster raw data using k-means and use sample data from each cluster to build a new datasetAnswer: BQUESTION 91A Data Scientist is training a multilayer perception (MLP) on a dataset with multiple classes. The target class of interest is unique compared to the other classes within the dataset, but it does not achieve and acceptable recall metric. The Data Scientist has already tried varying the number and size of the MLP's hidden layers, which has not significantly improved the results. A solution to improve recall must be implemented as quickly as possible.Which techniques should be used to meet these requirements?A.    Gather more data using Amazon Mechanical Turk and then retrainB.    Train an anomaly detection model instead of an MLPC.    Train an XGBoost model instead of an MLPD.    Add class weights to the MLP's loss function and then retrainAnswer: CQUESTION 92A Machine Learning Specialist works for a credit card processing company and needs to predict which transactions may be fraudulent in near-real time. Specifically, the Specialist must train a model that returns the probability that a given transaction may fraudulent.How should the Specialist frame this business problem?A. Streaming classificationB.    Binary classificationC.    Multi-category classificationD.    Regression classificationAnswer: CQUESTION 93A real estate company wants to create a machine learning model for predicting housing prices based on a historical dataset. The dataset contains 32 features.Which model will meet the business requirement?A.    Logistic regressionB.    Linear regressionC.    K-meansD.    Principal component analysis (PCA)Answer: BQUESTION 94A Machine Learning Specialist is applying a linear least squares regression model to a dataset with 1,000 records and 50 features. Prior to training, the ML Specialist notices that two features are perfectly linearly dependent.Why could this be an issue for the linear least squares regression model?A.   It could cause the backpropagation algorithm to fail during trainingB.    It could create a singular matrix during optimization, which fails to define a unique solutionC.    It could modify the loss function during optimization, causing it to fail during trainingD.    It could introduce non-linear dependencies within the data, which could invalidate the linear assumptions of the modelAnswer: CQUESTION 95Given the following confusion matrix for a movie classification model, what is the true class frequency for Romance and the predicted class frequency for Adventure?
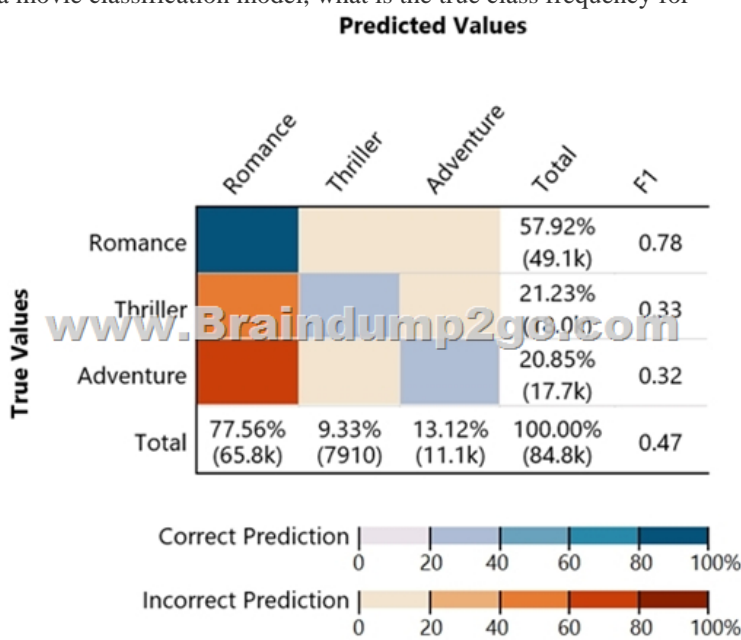


A.    The true class frequency for Romance is 77.56% and the predicted class frequency for Adventure is 20.85%B.    The true class

frequency for Romance is 57.92% and the predicted class frequency for Adventure is 13.12%C.    The true class frequency for Romance is 0.78% and the predicted class frequency for Adventure is (0.47-0.32)D.    The true class frequency for Romance is 77.56% * 0.78 and the predicted class frequency for Adventure is 20.85%*0.32Answer: BQUESTION 96A Machine Learning Specialist wants to bring a custom algorithm to Amazon SageMaker. The Specialist implements the algorithm in a Docker container supported by Amazon SageMaker.How should the Specialist package the Docker container so that Amazon SageMaker can launch the training correctly?A.    Modify the bash_profile file in the container and add a bash command to start the training programB.    Use CMD config in the Dockerfile to add the training program as a CMD of the imageC.    Configure the training program as an ENTRYPOINT named trainD.    Copy the training program to directory /opt/ml/trainAnswer: BQUESTION 97A Data Scientist needs to analyze employment data. The dataset contains approximately 10 million observations on people across 10 different features. During the preliminary analysis, the Data Scientist notices that income and age distributions are not normal. While income levels shows a right skew as expected, with fewer individuals having a higher income, the age distribution also shows a right skew, with fewer older individuals participating in the workforce.Which feature transformations can the Data Scientist apply to fix the incorrectly skewed data? (Choose two.)A.    Cross-validationB.    Numerical value binningC.    High-degree polynomial transformationD.    Logarithmic transformationE.    One hot encodingAnswer: ABQUESTION 98A web-based company wants to improve its conversion rate on its landing page. Using a large historical dataset of customer visits, the company has repeatedly trained a multi-class deep learning network algorithm on Amazon SageMaker. However, there is an overfitting problem: training data shows 90% accuracy in predictions, while test data shows 70% accuracy only.The company needs to boost the generalization of its model before deploying it into production to maximize conversions of visits to purchases.Which action is recommended to provide the HIGHEST accuracy model for the company's test and validation data?A.    Increase the randomization of training data in the mini-batches used in trainingB.    Allocate a higher proportion of the overall data to the training datasetC.    Apply L1 or L2 regularization and dropouts to the trainingD.    Reduce the number of layers and units (or neurons) from the deep learning networkAnswer: DQUESTION 99A Machine Learning Specialist is given a structured dataset on the shopping habits of a company's customer base. The dataset contains thousands of columns of data and hundreds of numerical columns for each customer. The Specialist wants to identify whether there are natural groupings for these columns across all customers and visualize the results as quickly as possible.What approach should the Specialist take to accomplish these tasks?A.    Embed the numerical features using the t-distributed stochastic neighbor embedding (t-SNE) algorithm and create a scatter plot.B.    Run k-means using the Euclidean distance measure for different values of k and create an elbow plot.C.    Embed the numerical features using the t-distributed stochastic neighbor embedding (t-SNE) algorithm and create a line graph.D.    Run k-means using the Euclidean distance measure for different values of k and create box plots for each numerical column within each cluster.Answer: BQUESTION 100A Machine Learning Specialist is planning to create a long-running Amazon EMR cluster. The EMR cluster will have 1 master node, 10 core nodes, and 20 task nodes. To save on costs, the Specialist will use Spot Instances in the EMR cluster.Which nodes should the Specialist launch on Spot Instances?A.    Master nodeB.    Any of the core nodesC.    Any of the task nodesD.    Both core and task nodesAnswer: AQUESTION 101A manufacturer of car engines collects data from cars as they are being driven. The data collected includes timestamp, engine temperature, rotations per minute (RPM), and other sensor readings. The company wants to predict when an engine is going to have a problem, so it can notify drivers in advance to get engine maintenance. The engine data is loaded into a data lake for training.Which is the MOST suitable predictive model that can be deployed into production?A.    Add labels over time to indicate which engine faults occur at what time in the future to turn this into a supervised learning problem. Use a recurrent neural network (RNN) to train the model to recognize when an engine might need maintenance for a certain fault.B.    This data requires an unsupervised learning algorithm. Use Amazon SageMaker k-means to cluster the data.C.    Add labels over time to indicate which engine faults occur at what time in the future to turn this into a supervised learning problem. Use a convolutional neural network (CNN) to train the model to recognize when an engine might need maintenance for a certain fault. D.    This data is already formulated as a time series. Use Amazon SageMaker seq2seq to model the time series.Answer: BQUESTION 102A company wants to predict the sale prices of houses based on available historical sales data. The target variable in the company's dataset is the sale price. The features include parameters such as the lot size, living area measurements, non-living area measurements, number of bedrooms, number of bathrooms, year built, and postal code. The company wants to use multi-variable linear regression to predict house sale prices.Which step should a machine learning specialist take to remove features that are irrelevant for the analysis and reduce the model's complexity?A.    Plot a histogram of the features and compute their standard deviation. Remove features with high variance.B.    Plot a histogram of the features and compute their standard deviation. Remove features with low variance.C.    Build a heatmap showing the correlation of the dataset against itself. Remove features with low mutual correlation scores.D.    Run a correlation check of all features against the target variable. Remove features with low target variable correlation scores.Answer: D

<u>Resources From:</u>1.2020 Latest Braindump2go MLS-C01 Exam Dumps (PDF & VCE) Free Share:
https://www.braindump2go.com/mls-c01.html2.2020 Latest Braindump2go MLS-C01 PDF and MLS-C01 VCE Dumps Free Share:
https://drive.google.com/drive/folders/1eX--L9LzE21hzqPIkigeo1QoAGNWL4vd?usp=sharing3.2020 Free Braindump2go
MLS-C01 PDF Download:https://www.braindump2go.com/free-online-pdf/MLS-C01-PDF(93-104).pdf
https://www.braindump2go.com/free-online-pdf/MLS-C01-PDF-Dumps(82-92).pdf
https://www.braindump2go.com/free-online-pdf/MLS-C01-VCE-Dumps(105-115).pdfFree Resources from Braindump2go,We
Devoted to Helping You 100% Pass All Exams!