

MLS-C01 Fragen & Antworten & MLS-C01 Studienführer & MLS-C01 Prüfungsvorbereitung



Außerdem sind jetzt einige Teile dieser ITZert MLS-C01 Prüfungsfragen kostenlos erhältlich: <https://drive.google.com/open?id=1keNn1E7Gr6HsvXBw4AQOkEfdqXvIV72p>

Die Schulungsunterlagen zur Amazon MLS-C01 Zertifizierungsprüfung von ITZert sind unvergleichbar. Das hat nicht nur mit der Qualität zu tun. Am wichtigsten ist es, dass Die Schulungsunterlagen zur Amazon MLS-C01 Zertifizierungsprüfung von ITZert mit allen IT-Zertifizierungen im Einklang sind. So kümmern sich viele Kandidaten um uns. Sie glauben in uns und sind von uns abhängig. Das hat genau unsere Stärke reflektiert. Sie werden sicher Ihren Freuden nach dem Kauf unserer Produkte ITZert empfehlen. Denn es kann Ihnen wirklich sehr helfen.

Die Erlangung der Amazon MLS-C01-Zertifizierung ist eine ausgezeichnete Möglichkeit für Fachleute, ihre Expertise in Machine Learning zu demonstrieren und ihre Karriere voranzutreiben. Es ist auch eine wertvolle Akkreditierung für Organisationen, die nach qualifizierten Fachleuten im Bereich Machine Learning suchen. Durch die Zertifizierung in Amazon MLS-C01 können Kandidaten ihre Hingabe an die aktuellen Trends und Technologien im schnelllebigen Bereich des Machine Learnings zeigen.

Die Zertifizierungsprüfung von Amazon MLS-C01 (AWS Certified Machine Learning-Specialty) ist eine begehrte Zertifizierung für Personen, die ihr Fachwissen im Bereich des maschinellen Lernens demonstrieren möchten. Diese Zertifizierung wird von Amazon Web Services (AWS) angeboten, einem der führenden Cloud -Computing -Anbieter der Welt. Die Prüfung wurde entwickelt, um die Fähigkeiten und das Wissen von Personen in Bereichen wie Datenerfassung, Datenvorverarbeitung, Modelltraining und Modellbereitstellung zu testen.

Die Amazon MLS -C01 (AWS Certified Machine Learning - Specialty) ist eine Zertifizierungsprüfung, die Einzelpersonen auf ihr Wissen und ihre Fähigkeiten im Zusammenhang mit maschinellem Lernen auf der Amazon Web Services (AWS) -Plattform testet. Diese Prüfung richtet sich an Personen, die zertifizierte Fachleute im Bereich des maschinellen Lernens werden möchten und ihr Fachwissen zum Entwerfen, Implementieren, Bereitstellen und Aufrechterhalten von Lösungen für maschinelles Lernen auf AWS präsentieren möchten.

>> MLS-C01 Online Praxisprüfung <<

MLS-C01 Zertifizierungsfragen & MLS-C01 PDF

Sie sollen Methode zum Erfolg, nicht Einwände für die Niederlage finden. Es ist doch nicht so schwer, die Amazon MLS-C01 Zertifizierungsprüfung zu bestehen. Die Schulungsunterlagen zur Amazon MLS-C01 Zertifizierungsprüfung von ITZert zu wählen ist eine gute Wahl, die Ihnen zum Bestehen der Amazon MLS-C01 Prüfung verhelfen. Sie sind auch die beste Abkürzung zum Erfolg. Jeder will Erfolg erlangen. Hauptsache, man muss richtige Wahl treffen.

Amazon AWS Certified Machine Learning - Specialty MLS-C01 Prüfungsfragen mit Lösungen (Q213-Q218):

213. Frage

A gaming company has launched an online game where people can start playing for free but they need to pay if they choose to use certain features. The company needs to build an automated system to predict whether or not a new user will become a paid user within 1 year. The company has gathered a labeled dataset from 1 million users. The training dataset consists of 1,000 positive samples (from users who ended up paying within 1 year) and 999,000 negative samples (from users who did not use any paid features). Each data sample consists of 200 features including user age, device, location, and play patterns. Using this dataset for training, the Data Science team trained a random forest model that converged with over 99% accuracy on the training set. However, the prediction results on a test dataset were not satisfactory. Which of the following approaches should the Data Science team take to mitigate this issue? (Select TWO.)

- A. Add more deep trees to the random forest to enable the model to learn more features.
- B. Change the cost function so that false positives have a higher impact on the cost value than false negatives.
- C. Generate more positive samples by duplicating the positive samples and adding a small amount of noise to the duplicated data.
- D. Change the cost function so that false negatives have a higher impact on the cost value than false positives.
- E. Indicate a copy of the samples in the test database in the training dataset.

Antwort: C,D

Begründung:

The Data Science team is facing a problem of imbalanced data, where the positive class (paid users) is much less frequent than the negative class (non-paid users). This can cause the random forest model to be biased towards the majority class and have poor performance on the minority class. To mitigate this issue, the Data Science team can try the following approaches:

C). Generate more positive samples by duplicating the positive samples and adding a small amount of noise to the duplicated data. This is a technique called data augmentation, which can help increase the size and diversity of the training data for the minority class. This can help the random forest model learn more features and patterns from the positive class and reduce the imbalance ratio.

D). Change the cost function so that false negatives have a higher impact on the cost value than false positives. This is a technique called cost-sensitive learning, which can assign different weights or costs to different classes or errors. By assigning a higher cost to false negatives (predicting non-paid when the user is actually paid), the random forest model can be more sensitive to the minority class and try to minimize the misclassification of the positive class.

Bagging and Random Forest for Imbalanced Classification

Surviving in a Random Forest with Imbalanced Datasets

machine learning - random forest for imbalanced data? - Cross Validated
Biased Random Forest For Dealing With the Class Imbalance Problem

214. Frage

A 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 line graph.
- B. Run k-means using the Euclidean distance measure for different values of k and create an elbow plot.
- C. Run k-means using the Euclidean distance measure for different values of k and create box plots for each numerical column within each cluster.
- D. Embed the numerical features using the t-distributed stochastic neighbor embedding (t-SNE) algorithm and create a scatter plot.

Antwort: D

Begründung:

The best approach to identify and visualize the natural groupings for the numerical columns across all customers is to embed the numerical features using the t-distributed stochastic neighbor embedding (t-SNE) algorithm and create a scatter plot. t-SNE is a dimensionality reduction technique that can project high-dimensional data into a lower-dimensional space, while preserving the local structure and distances of the data points. A scatter plot can then show the clusters of data points in the reduced space, where each point represents a customer and the color indicates the cluster membership. This approach can help the Specialist quickly explore the patterns and similarities among the customers based on their numerical features.

The other options are not as effective or efficient as the t-SNE approach. Running k-means for different values of k and creating an elbow plot can help determine the optimal number of clusters, but it does not provide a visual representation of the clusters or the customers. Embedding the numerical features using t-SNE and creating a line graph does not make sense, as a line graph is used to show the change of a variable over time, not the distribution of data points in a space. Running k-means for different values of k and creating box plots for each numerical column within each cluster can provide some insights into the statistics of each cluster, but it is very time-consuming and cumbersome to create and compare thousands of box plots.

References:

* Dimensionality Reduction - Amazon SageMaker

* Visualize high dimensional data using t-SNE - Amazon SageMaker

215. Frage

A gaming company has launched an online game where people can start playing for free but they need to pay if they choose to use certain features. The company needs to build an automated system to predict whether or not a new user will become a paid user within 1 year. The company has gathered a labeled dataset from 1 million users. The training dataset consists of 1,000 positive samples (from users who ended up paying within 1 year) and 999,000 negative samples (from users who did not use any paid features). Each data sample consists of 200 features including user age, device, location, and play patterns. Using this dataset for training, the Data Science team trained a random forest model that converged with over 99% accuracy on the training set. However, the prediction results on a test dataset were not satisfactory.

Which of the following approaches should the Data Science team take to mitigate this issue? (Select TWO.)

- A. Add more deep trees to the random forest to enable the model to learn more features.
- B. Change the cost function so that false positives have a higher impact on the cost value than false negatives.
- C. Generate more positive samples by duplicating the positive samples and adding a small amount of noise to the duplicated data.
- D. Change the cost function so that false negatives have a higher impact on the cost value than false positives.
- E. Indicate a copy of the samples in the test database in the training dataset.

Antwort: C,D

216. Frage

A company is using Amazon Textract to extract textual data from thousands of scanned text-heavy legal documents daily. The company uses this information to process loan applications automatically. Some of the documents fail business validation and are returned to human reviewers, who investigate the errors. This activity increases the time to process the loan applications.

What should the company do to reduce the processing time of loan applications?

- A. Use Amazon Rekognition's feature to detect text in an image to extract the data from scanned images. Use this information to process the loan applications.
- B. Configure Amazon Textract to route low-confidence predictions to Amazon Augmented AI (Amazon A2I). Perform a manual review on those words before performing a business validation.
- C. Use an Amazon Textract synchronous operation instead of an asynchronous operation.
- D. Configure Amazon Textract to route low-confidence predictions to Amazon SageMaker Ground Truth. Perform a manual review on those words before performing a business validation.

Antwort: B

Begründung:

The company should configure Amazon Textract to route low-confidence predictions to Amazon Augmented AI (Amazon A2I). Amazon A2I is a service that allows you to implement human review of machine learning (ML) predictions. It also comes integrated with some of the Artificial Intelligence (AI) services such as Amazon Textract. By using Amazon A2I, the company can perform a manual review on those words that have low confidence scores before performing a business validation. This will help reduce the

processing time of loan applications by avoiding errors and rework.

Option A is incorrect because Amazon SageMaker Ground Truth is not a suitable service for human review of Amazon Textract predictions. Amazon SageMaker Ground Truth is a service that helps you build highly accurate training datasets for machine learning. It allows you to label your own data or use a workforce of human labelers. However, it does not provide an easy way to integrate with Amazon Textract and route low- confidence predictions for human review.

Option B is incorrect because using an Amazon Textract synchronous operation instead of an asynchronous operation will not reduce the processing time of loan applications. A synchronous operation is a request- response operation that returns the results immediately. An asynchronous operation is a start-and-check operation that returns a job identifier that you can use to check the status and results later. The choice of operation depends on the size and complexity of the document, not on the confidence of the predictions.

Option D is incorrect because using Amazon Rekognition's feature to detect text in an image to extract the data from scanned images is not a better alternative than using Amazon Textract. Amazon Rekognition is a service that provides computer vision capabilities, such as face recognition, object detection, and scene analysis. It can also detect text in an image, but it does not provide the same level of accuracy and functionality as Amazon Textract. Amazon Textract can not only detect text, but also extract data from tables and forms, and understand the layout and structure of the document.

Amazon Augmented AI

Amazon SageMaker Ground Truth

Amazon Textract Operations

Amazon Rekognition

217. Frage

A company has an ecommerce website with a product recommendation engine built in TensorFlow. The recommendation engine endpoint is hosted by Amazon SageMaker. Three compute-optimized instances support the expected peak load of the website. Response times on the product recommendation page are increasing at the beginning of each month. Some users are encountering errors. The website receives the majority of its traffic between 8 AM and 6 PM on weekdays in a single time zone.

Which of the following options are the MOST effective in solving the issue while keeping costs to a minimum? (Choose two.)

- A. Create a new endpoint configuration with two production variants.
- **B. Configure the endpoint to use Amazon Elastic Inference (EI) accelerators.**
- C. Reconfigure the endpoint to use burstable instances.
- D. Deploy a second instance pool to support a blue/green deployment of models.
- **E. Configure the endpoint to automatically scale with the Invocations Per Instance metric.**

Antwort: B,E

Begründung:

Explanation

The solution A and C are the most effective in solving the issue while keeping costs to a minimum. The solution A and C involve the following steps:

Configure the endpoint to use Amazon Elastic Inference (EI) accelerators. This will enable the company to reduce the cost and latency of running TensorFlow inference on SageMaker. Amazon EI provides GPU-powered acceleration for deep learning models without requiring the use of GPU instances. Amazon EI can attach to any SageMaker instance type and provide the right amount of acceleration based on the workload¹.

Configure the endpoint to automatically scale with the Invocations Per Instance metric. This will enable the company to adjust the number of instances based on the demand and traffic patterns of the website.

The Invocations Per Instance metric measures the average number of requests that each instance processes over a period of time.

By using this metric, the company can scale out the endpoint when the load increases and scale in when the load decreases. This can improve the response time and availability of the product recommendation engine².

The other options are not suitable because:

Option B: Creating a new endpoint configuration with two production variants will not solve the issue of increasing response time and errors. Production variants are used to split the traffic between different models or versions of the same model. They can be useful for testing, updating, or A/B testing models. However, they do not provide any scaling or acceleration benefits for the inference workload³.

Option D: Deploying a second instance pool to support a blue/green deployment of models will not solve the issue of increasing response time and errors. Blue/green deployment is a technique for updating models without downtime or disruption. It involves creating a new endpoint configuration with a different instance pool and model version, and then shifting the traffic from the old endpoint to the new endpoint gradually. However, this technique does not provide any scaling or acceleration benefits for the inference workload⁴.

Option E: Reconfiguring the endpoint to use burstable instances will not solve the issue of increasing response time and errors.

Burstable instances are instances that provide a baseline level of CPU performance with the ability to burst above the baseline when

References:

- ### 218. Frage

• • • • •

MLS-C01 Zertifizierungsfragen: https://www.itcert.com/MLS-C01_valid-braindumps.html

- P.S. Kostenlose und neue MLS-C01 Prüfungsfragen sind auf Google Drive freigegeben von ITZert verfügbar: <https://drive.google.com/open?id=1keNn1E7Gr6HsvXBw4AQOOkEfdqXvIv72p>