ExamsIndex
Practice Materials & Tests

DEMO VERSION

Amazon

AIF-C01 Exam

AWS Certified AI Practitioner Exam

Exam Latest Version: 18.1

Question 1. (Single Select)

An AI practitioner trained a custom model on Amazon Bedrock by using a training dataset that contains confidential data. The AI practitioner wants to ensure that the custom model does not generate inference responses based on confidential data.

How should the AI practitioner prevent responses based on confidential data?

A: Delete the custom model. Remove the confidential data from the training dataset. Retrain the custom model.
B: Mask the confidential data in the inference responses by using dynamic data masking.
C: Encrypt the confidential data in the inference responses by using Amazon SageMaker.
D: Encrypt the confidential data in the custom model by using AWS Key Management Service (AWS KMS).

Correct Answer: A

Explanation:

When a model is trained on a dataset containing confidential or sensitive data, the model may inadvertently learn patterns from this data, which could then be reflected in its inference responses. To ensure that a model does not generate responses based on confidential data, the most effective approach is to remove the confidential data from the training dataset and then retrain the model.

Explanation of Each Option:

Option A (Correct): "Delete the custom model. Remove the confidential data from the training dataset. Retrain the custom model."This option is correct because it directly addresses the core issue: the model has been trained on confidential data. The only way to ensure that the model does not produce inferences based on this data is to remove the confidential information from the training dataset and then retrain the model from scratch. Simply deleting the model and retraining it ensures that no confidential data is learned or retained by the model. This approach follows the best practices recommended by AWS for handling sensitive data when using machine learning services like Amazon Bedrock.

Option B: "Mask the confidential data in the inference responses by using dynamic data masking."This option is incorrect because dynamic data masking is typically used to mask or obfuscate sensitive data in a database. It does not address the core problem of the model

beingtrained on confidential data. Masking data in inference responses does not prevent the model from using confidential data it learned during training.

Option C: "Encrypt the confidential data in the inference responses by using Amazon SageMaker."This option is incorrect because encrypting the inference responses does not prevent the model from generating outputs based on confidential data. Encryption only secures the data at rest or in transit but does not affect the model's underlying knowledge or training process.

Option D: "Encrypt the confidential data in the custom model by using AWS Key Management Service (AWS KMS)."This option is incorrect as well because encrypting the data within the model does not prevent the model from generating responses based on the confidential data it learned during training. AWS KMS can encrypt data, but it does not modify the learning that the model has already performed.

AWS AI Practitioner Reference:

Data Handling Best Practices in AWS Machine Learning: AWS advises practitioners to carefully handle training data, especially when it involves sensitive or confidential information. This includes preprocessing steps like data anonymization or removal of sensitive data before using it to train machine learning models.

Amazon Bedrock and Model Training Security: Amazon Bedrock provides foundational models and customization capabilities, but any training involving sensitive data should follow best practices, such as removing or anonymizing confidential data to prevent unintended data leakage.

Question 2. (Multi Select)

A loan company is building a generative AI-based solution to offer new applicants discounts based on specific business criteri

a. The company wants to build and use an AI model responsibly to minimize bias that could negatively affect some customers.

Which actions should the company take to meet these requirements? (Select TWO.)

A: Detect imbalances or disparities in the data.
B: Ensure that the model runs frequently.
C: Evaluate the model's behavior so that the company can provide transparency to stakeholders.
D: Use the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) technique to ensure that the model is 100% accurate.
E: Ensure that the model's inference time is within the accepted limits.

Correct Answer: A, C

Explanation:

To build an AI model responsibly and minimize bias, it is essential to ensure fairness and transparency throughout the model development and deployment process. This involves detecting and mitigating data imbalances and thoroughly evaluating the model's behavior to understand its impact on different groups.

Option A (Correct): "Detect imbalances or disparities in the data": This is correct because identifying and addressing data imbalances or disparities is a critical step in reducing bias. AWS provides tools like Amazon SageMaker Clarify to detect bias during data preprocessing and model training.

Option C (Correct): "Evaluate the model's behavior so that the company can provide transparency to stakeholders": This is correct because evaluating the model's behavior for fairness and accuracy is key to ensuring that stakeholders understand how the model makes decisions. Transparency is a crucial aspect of responsible AI.

Option B: "Ensure that the model runs frequently" is incorrect because the frequency of model runs does not address bias.

Option D: "Use the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) technique to ensure that the model is 100% accurate" is incorrect because ROUGE is a metric for evaluating the quality of text summarization models, not for minimizing bias.

Option E: "Ensure that the model's inference time is within the accepted limits" is incorrect as it relates to performance, not bias reduction.

AWS AI Practitioner Reference:

Amazon SageMaker Clarify: AWS offers tools such as SageMaker Clarify for detecting bias in datasets and models, and for understanding model behavior to ensure fairness and transparency.

Responsible AI Practices: AWS promotes responsible AI by advocating for fairness, transparency, and inclusivity in model development and deployment.

Question 3. (HOTSPOT)

A company wants to build an ML application.

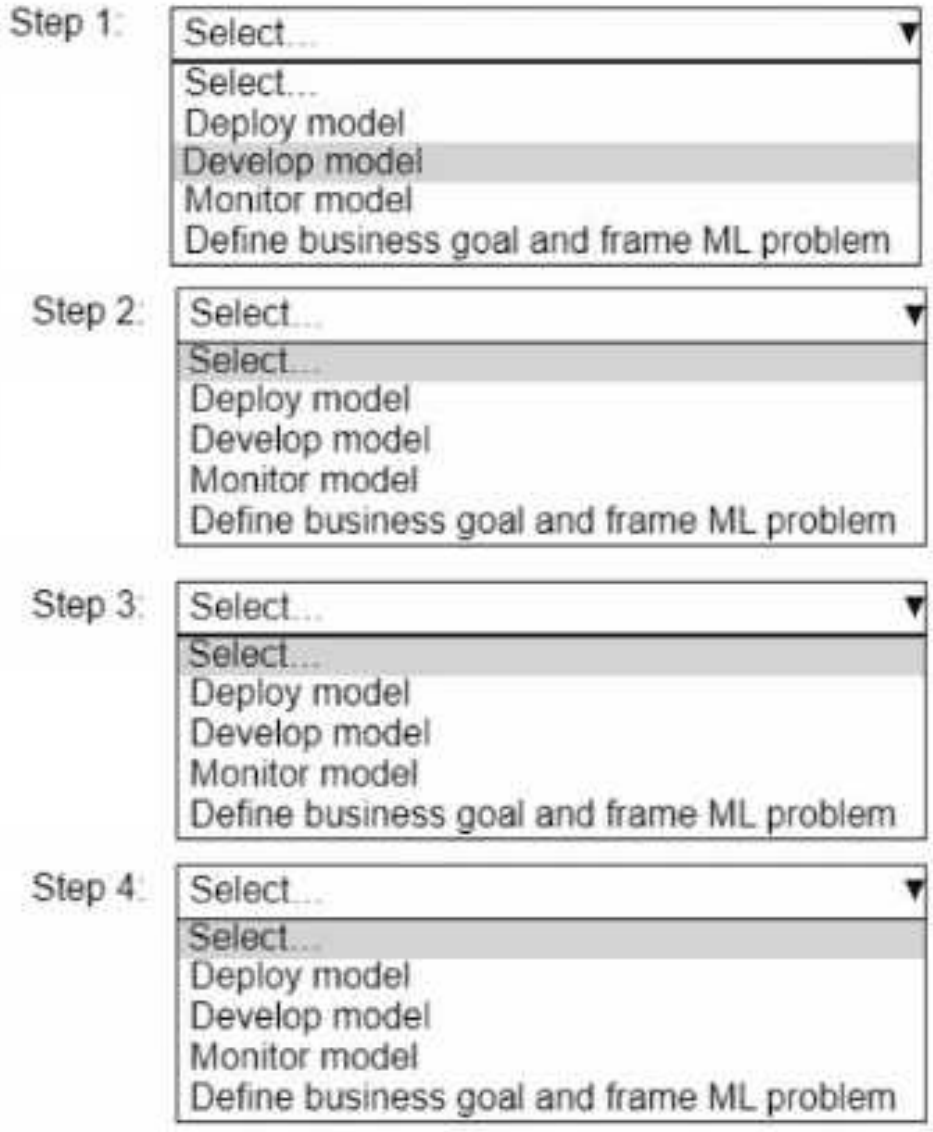
Select and order the correct steps from the following list to develop a well-architected ML workload. Each step should be selected one time. (Select and order FOUR.)

• Deploy model

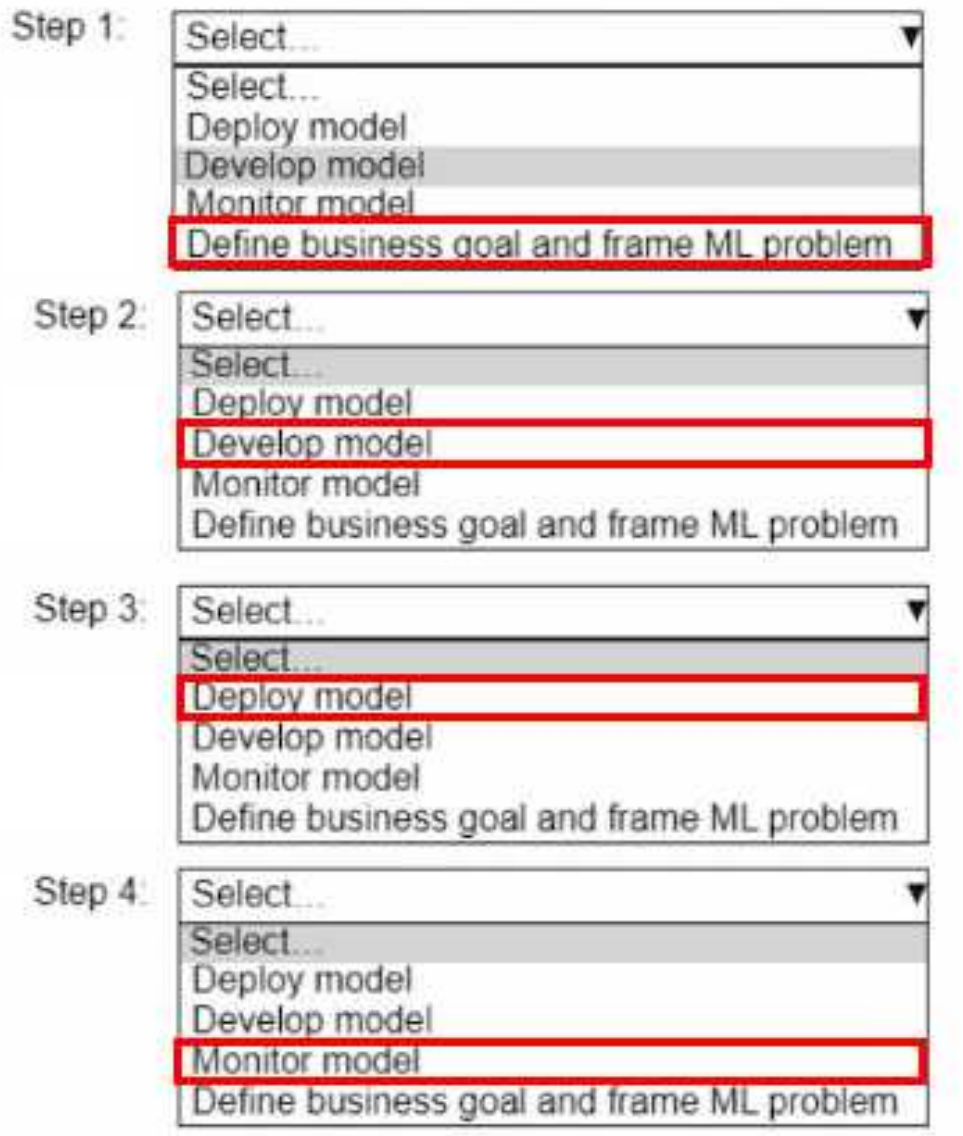
• Develop model

• Monitor model

• Define business goal and frame ML problem



Correct Answer:



Question 4. (Single Select)

A company has petabytes of unlabeled customer data to use for an advertisement campaign. The company wants to classify its customers into tiers to advertise and promote the company's products.

Which methodology should the company use to meet these requirements?

A: Supervised learning

B: Unsupervised learning
C: Reinforcement learning
D: Reinforcement learning from human feedback (RLHF)

Correct Answer: B

Explanation:

Unsupervised learning is the correct methodology for classifying customers into tiers when the data is unlabeled, as it does not require predefined labels or outputs.

Unsupervised Learning:

This type of machine learning is used when the data has no labels or pre-defined categories. The goal is to identify patterns, clusters, or associations within the data.

In this case, the company has petabytes of unlabeled customer data and needs to classify customers into different tiers. Unsupervised learning techniques like clustering (e.g., K-Means, Hierarchical Clustering) can group similar customers based on various attributes without any prior knowledge or labels.

Why Option B is Correct:

Handling Unlabeled Data: Unsupervised learning is specifically designed to work with unlabeled data, making it ideal for the company’s need to classify customer data.

Customer Segmentation: Techniques in unsupervised learning can be used to find natural groupings within customer data, such as identifying high-value vs. low-value customers or segmenting based on purchasing behavior.

Why Other Options are Incorrect:

A . Supervised learning: Requires labeled data with input-output pairs to train the model, which is not suitable since the company's data is unlabeled.

C . Reinforcement learning: Focuses on training an agent to make decisions by maximizing some notion of cumulative reward, which does not align with the company's need for customer classification.

D . Reinforcement learning from human feedback (RLHF): Similar to reinforcement learning but involves human feedback to refine the model’s behavior; it is also not appropriate for classifying unlabeled customer data.

Question 5. (Single Select)

A company makes forecasts each quarter to decide how to optimize operations to meet expected demand. The company uses ML models to make these forecasts.

An AI practitioner is writing a report about the trained ML models to provide transparency and explainability to company stakeholders.

What should the AI practitioner include in the report to meet the transparency and explainability requirements?

A: Code for model training
B: Partial dependence plots (PDPs)
C: Sample data for training
D: Model convergence tables

Correct Answer: B

Explanation:

Partial dependence plots (PDPs) are visual tools used to show the relationship between a feature (or a set of features) in the data and the predicted outcome of a machine learning model. They are highly effective for providing transparency and explainability of the model's behavior to stakeholders by illustrating how different input variables impact the model's predictions.

Option B (Correct): "Partial dependence plots (PDPs)": This is the correct answer because PDPs help to interpret how the model's predictions change with varying values of input features, providing stakeholders with a clearer understanding of the model's decision-making process.

Option A: "Code for model training" is incorrect because providing the raw code for model training may not offer transparency or explainability to non-technical stakeholders.

Option C: "Sample data for training" is incorrect as sample data alone does not explain how the model works or its decision-making process.

Option D: "Model convergence tables" is incorrect. While convergence tables can show the training process, they do not provide insights into how input features affect the model's predictions.

AWS AI Practitioner Reference:

Explainability in AWS Machine Learning: AWS provides various tools for model explainability, such as Amazon SageMaker Clarify, which includes PDPs to help explain the impact of different features on the model's predictions.

ExamsIndex

Demo PDF Complete

Your AIF-C01 Demo (5 Questions)

Get the Complete Version

Full Questions with Detailed Explanations

Interactive Web-Based Exams Available

To get 30% off, use Coupon Code: NEWYEAR30

https://examsindex.com/exam/aif-c01