Amazon.AIF-C01.v2025-08-07.q87

Exam Code:	AIF-C01
Exam Name:	AWS Certified AI Practitioner
Certification Provider:	Amazon
Free Question Number:	87
Version:	v2025-08-07
# of views:	143
# of Questions views:	870
https://www.freeqas.com/qa/Amazon/AIF-C01/Amazon.AIF-C01.v2025-08-07.q87.html	

NEW QUESTION: 1
A digital devices company wants to predict customer demand for memory hardware. The company does not have coding experience or knowledge of ML algorithms and needs to develop a data-driven predictive model. The company needs to perform analysis on internal data and external data.
Which solution will meet these requirements?
A. Store the data in Amazon S3. Create ML models and demand forecast predictions by using Amazon SageMaker built-in algorithms that use the data from Amazon S3.
B. Import the data into Amazon SageMaker Data Wrangler. Create ML models and demand forecast predictions by using SageMaker built-in algorithms.
C. Import the data into Amazon SageMaker Data Wrangler. Build ML models and demand forecast predictions by using an Amazon Personalize Trending-Now recipe.
D. Import the data into Amazon SageMaker Canvas. Build ML models and demand forecast predictions by selecting the values in the data from SageMaker Canvas.
Answer: D (LEAVE A REPLY)
I'll continue to format the rest. Let me know if you would like me to provide them all in one go or in parts.

NEW QUESTION: 2
A company wants to develop an AI application to help its employees check open customer claims, identify details for a specific claim, and access documents for a claim. Which solution meets these requirements?
A. Use Agents for Amazon Bedrock with Amazon Fraud Detector to build the application.
B. Use Agents for Amazon Bedrock with Amazon Bedrock knowledge bases to build the application.
C. Use Amazon Personalize with Amazon Bedrock knowledge bases to build the application.

D. Use Amazon SageMaker AI to build the application by training a new ML model.

Answer: B (LEAVE A REPLY)

The company wants an AI application to help employees check open customer claims, identify claim details, and access related documents. Agents for Amazon Bedrock can automate tasks by interacting with external systems, while Amazon Bedrock knowledge bases provide a repository of information (e.g., claim details and documents) that the agent can query to respond to employee requests, making this the best solution.

Exact Extract from AWS AI Documents:

From the AWS Bedrock User Guide:

"Agents for Amazon Bedrock enable developers to build applications that can perform tasks by interacting with external systems and data sources. When paired with Amazon Bedrock knowledge bases, agents can access structured and unstructured data, such as documents or databases, to provide detailed responses for use cases like customer service or claims management." (Source: AWS Bedrock User Guide, Agents and Knowledge Bases) Detailed Explanation:

Option A: Use Agents for Amazon Bedrock with Amazon Fraud Detector to build the application.Amazon Fraud Detector is for detecting fraudulent activities, not for managing customer claims or accessing documents. This option is irrelevant.

Option B: Use Agents for Amazon Bedrock with Amazon Bedrock knowledge bases to build the application.

This is the correct answer. Agents for Amazon Bedrock can interact with knowledge bases to retrieve claim details and documents, enabling employees to check open claims and access relevant information.

Option C: Use Amazon Personalize with Amazon Bedrock knowledge bases to build the application.Amazon Personalize is for building recommendation systems, not for retrieving claim details or documents. This option does not meet the requirements.

Option D: Use Amazon SageMaker AI to build the application by training a new ML model.Training a new ML model on SageMaker is unnecessary and complex for this use case, as the task can be efficiently handled by Agents and knowledge bases on Amazon Bedrock.

References:

AWS Bedrock User Guide: Agents and Knowledge Bases (https://docs.aws.amazon.com/bedrock/latest /userguide/agents.html)

AWS AI Practitioner Learning Path: Module on Generative AI and Knowledge Bases

Amazon Bedrock Developer Guide: Building AI Applications (https://aws.amazon.com/bedrock/)

NEW QUESTION: 3

A company is using a large language model (LLM) on Amazon Bedrock to build a chatbot. The chatbot processes customer support requests. To resolve a request, the customer and the chatbot must interact a few times.
Which solution gives the LLM the ability to use content from previous customer messages?
A. Turn on model invocation logging to collect messages.
B. Add messages to the model prompt.
C. Use Amazon Personalize to save conversation history.
D. Use Provisioned Throughput for the LLM.
Answer: B (LEAVE A REPLY)
The company is building a chatbot using an LLM on Amazon Bedrock, and the chatbot needs to use content from previous customer messages to resolve requests. Adding previous messages to the model prompt (also known as providing conversation history) enables the LLM to maintain context across interactions, allowing it to respond coherently based on the ongoing conversation.
Exact Extract from AWS AI Documents:
From the AWS Bedrock User Guide:
"To enable a large language model (LLM) to maintain context in a conversation, you can include previous messages in the model prompt. This approach, often referred to as providing conversation history, allows the LLM to generate responses that are contextually relevant toprior interactions." (Source: AWS Bedrock User Guide, Building Conversational Applications) Detailed Explanation:
* Option A: Turn on model invocation logging to collect messages.Model invocation logging records interactions for auditing or debugging but does not provide the LLM with access to previous messages during inference to maintain conversation context.
* Option B: Add messages to the model prompt.This is the correct answer. Including previous messages in the prompt gives the LLM the conversation history it needs to respond appropriately, a common practice for chatbots on Amazon Bedrock.
* Option C: Use Amazon Personalize to save conversation history.Amazon Personalize is for building recommendation systems, not for managing conversation history in a chatbot. This option is irrelevant.
* Option D: Use Provisioned Throughput for the LLM.Provisioned Throughput in Amazon Bedrock ensures consistent performance for model inference but does not address the need to use previous messages in the conversation.
References:
AWS Bedrock User Guide: Building Conversational Applications (https://docs.aws.amazon.com/bedrock/latest/userguide/conversational-apps.html)
AWS AI Practitioner Learning Path: Module on Generative AI and Chatbots Amazon Bedrock Developer Guide: Managing Conversation Context (https://aws.amazon.com/bedrock/)

NEW QUESTION: 4
A company is implementing the Amazon Titan foundation model (FM) by using Amazon Bedrock. The company needs to supplement the model by using relevant data from the company's private data sources.
Which solution will meet this requirement?
A. Choose a lower temperature value
B. Enable model invocation logging
C. Use a different FM
D. Create an Amazon Bedrock knowledge base
Answer: D (LEAVE A REPLY)

NEW QUESTION: 5
A company has built an image classification model to predict plant diseases from photos of plant leaves. The company wants to evaluate how many images the model classified correctly.
Which evaluation metric should the company use to measure the model's performance?
A. Accuracy
B. Root mean squared error (RMSE)
C. R-squared score
D. Learning rate
Answer: (SHOW ANSWER)

NEW QUESTION: 6
A loan company is building a generative AI-based solution to offer new applicants discounts based on specific business criteria. 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.
Answer: A,C (LEAVE A REPLY)
To build and use an AI model responsibly, especially in sensitive applications like loan approvals, it's crucial to address potential biases and ensure transparency:
* Detect imbalances or disparities in the data (Option A): Analyzing the training data for imbalances or disparities is essential. Imbalanced data can lead to models that are biased towards the majority class, potentially disadvantaging certain groups. By identifying and mitigating these imbalances, the company can reduce the risk of biased predictions.

* Evaluate the model's behavior to provide transparency to stakeholders (Option C): Regularly assessing the model's outputs and decision-making processes allows the company to understand how decisions are made. This evaluation fosters transparency, enabling the company to explain model behavior to stakeholders and ensure that the model operates as intended without unintended biases.
Options B, D, and E, while relevant to model performance and evaluation, do not directly address the responsible use of AI concerning bias and transparency.

NEW QUESTION: 7
A security company is using Amazon Bedrock to run foundation models (FMs). The company wants to ensure that only authorized users invoke the models. The company needs to identify any unauthorized access attempts to set appropriate AWS Identity and Access Management (IAM) policies and roles for future iterations of the FMs.
Which AWS service should the company use to identify unauthorized users that are trying to access Amazon Bedrock?
A. AWS CloudTrail
B. AWS Trusted Advisor
C. Amazon Fraud Detector
D. AWS Audit Manager
Answer: (SHOW ANSWER)

NEW QUESTION: 8
How can companies use large language models (LLMs) securely on Amazon Bedrock?
A. Design clear and specific prompts. Configure AWS Identity and Access Management (IAM) roles and policies by using least privilege access.
B. Use Amazon CloudWatch Logs to make models explainable and to monitor for bias.
C. Enable AWS Audit Manager for automatic model evaluation jobs.
D. Enable Amazon Bedrock automatic model evaluation jobs.
Answer: A (LEAVE A REPLY)

NEW QUESTION: 9
A company wants to use generative AI to increase developer productivity and software development. The company wants to use Amazon Q Developer.
What can Amazon Q Developer do to help the company meet these requirements?
A. Run an application without provisioning or managing servers.
B. Enable voice commands for coding and providing natural language search.
C. Convert audio files to text documents by using ML models.
D. Create software snippets, reference tracking, and open-source license tracking.
Answer: B (LEAVE A REPLY)

NEW QUESTION: 10

A company is using a pre-trained large language model (LLM) to build a chatbot for product recommendations. The company needs the LLM outputs to be short and written in a specific language.
Which solution will align the LLM response quality with the company's expectations?
A. Adjust the prompt.
B. Choose an LLM of a different size.
C. Increase the temperature.
D. Increase the Top K value.
Answer: A (LEAVE A REPLY)
Adjusting the prompt is the correct solution to align the LLM outputs with the company's expectations for short, specific language responses.
* Adjust the Prompt:
* Modifying the prompt can guide the LLM to produce outputs that are shorter and tailored to the desired language.
* A well-crafted prompt can provide specific instructions to the model, such as "Answer in a short sentence in Spanish."
* Why Option A is Correct:
* Control Over Output: Adjusting the prompt allows for direct control over the style, length, and language of the LLM outputs.
* Flexibility: Prompt engineering is a flexible approach to refining the model's behavior without modifying the model itself.
* Why Other Options are Incorrect:
* B. Choose an LLM of a different size: The model size does not directly impact the response length or language.
* C. Increase the temperature: Increases randomness in responses but does not ensure brevity or specific language.
* D. Increase the Top K value: Affects diversity in model output but does not align directly with response length or language specificity.

NEW QUESTION: 11
A company wants to use a large language model (LLM) on Amazon Bedrock for sentiment analysis. The company wants to classify the sentiment of text passages as positive or negative.
Which prompt engineering strategy meets these requirements?
A. Provide examples of text passages with corresponding positive or negative labels in the prompt followed by the new text passage to be classified.
B. Provide a detailed explanation of sentiment analysis and how LLMs work in the prompt.
C. Provide the new text passage to be classified without any additional context or examples.
D. Provide the new text passage with a few examples of unrelated tasks, such as text summarization or question answering.

Answer: (SHOW ANSWER)
Providing examples of text passages with corresponding positive or negative labels in the prompt followed by the new text passage to be classified is the correct prompt engineering strategy for using a large language model (LLM) on Amazon Bedrock for sentiment analysis.
* Example-Driven Prompts:
* This strategy, known as few-shot learning, involves giving the model examples of input-output pairs (e.g., text passages with their sentiment labels) to help it understand the task context.
* It allows the model to learn from these examples and apply the learned pattern to classify new text passages correctly.
* Why Option A is Correct:
* Guides the Model: Providing labeled examples teaches the model how to perform sentiment analysis effectively, increasing accuracy.
* Contextual Relevance: Aligns the model's responses to the specific task of classifying sentiment.
* Why Other Options are Incorrect:
* B. Detailed explanation of sentiment analysis: Unnecessary for the model's operation; it requires examples, not explanations.
* C. New text passage without context: Provides no guidance or learning context for the model.
* D. Unrelated task examples: Would confuse the model and lead to inaccurate results.

NEW QUESTION: 12
A company has built an image classification model to predict plant diseases from photos of plant leaves. The company wants to evaluate how many images the model classified correctly.
Which evaluation metric should the company use to measure the model's performance?
A. R-squared score
B. Accuracy
C. Root mean squared error (RMSE)
D. Learning rate
Answer: (SHOW ANSWER)
Accuracy is the most appropriate metric to measure the performance of an image classification model. It indicates the percentage of correctly classified images out of the total number of images. In the context of classifying plant diseases from images, accuracy will help the company determine how well the model is performing by showing how many images were correctly classified.
* Option B (Correct): "Accuracy": This is the correct answer because accuracy measures the proportion of correct predictions made by the model, which is suitable for evaluating the performance of a classification model.

* Option A: "R-squared score" is incorrect as it is used for regression analysis, not classification tasks.
* Option C: "Root mean squared error (RMSE)" is incorrect because it is also used for regression tasks to measure prediction errors, not for classification accuracy.
* Option D: "Learning rate" is incorrect as it is a hyperparameter for training, not a performance metric.
AWS AI Practitioner References:
* Evaluating Machine Learning Models on AWS: AWS documentation emphasizes the use of appropriate metrics, like accuracy, for classification tasks.

NEW QUESTION: 13
A company wants to use language models to create an application for inference on edge devices. The inference must have the lowest latency possible.
Which solution will meet these requirements?
A. Deploy optimized large language models (LLMs) on edge devices.
B. Deploy optimized small language models (SLMs) on edge devices.
C. Incorporate a centralized small language model (SLM) API for asynchronous communication with edge devices.
D. Incorporate a centralized large language model (LLM) API for asynchronous communication with edge devices.
Answer: (SHOW ANSWER)

NEW QUESTION: 14
An ecommerce company is deploying a chatbot. The chatbot will give users the ability to ask questions about the company's products and receive details on users' orders. The company must implement safeguards for the chatbot to filter harmful content from the input prompts and chatbot responses.
Which AWS feature or resource meets these requirements?
A. Amazon Bedrock Guardrails
B. Amazon Bedrock Agents
C. Amazon Bedrock inference APIs
D. Amazon Bedrock custom models
Answer: A (LEAVE A REPLY)
The ecommerce company is deploying a chatbot that needs safeguards to filter harmful content from input prompts and responses. Amazon Bedrock Guardrails provide mechanisms to ensure responsible AI usage by filtering harmful content, such as hate speech, violence, or misinformation, making it the appropriate feature for this requirement.
Exact Extract from AWS AI Documents:
From the AWS Bedrock User Guide:
"Amazon Bedrock Guardrails enable developers to implement safeguards for generative AI applications, such as chatbots, by filtering harmful content in input prompts and model

responses. Guardrails include content filters, word filters, and denied topics to ensure safe and responsible outputs." (Source: AWS Bedrock User Guide, Guardrails for Responsible AI) Detailed Explanation:
* Option A: Amazon Bedrock GuardrailsThis is the correct answer. Amazon Bedrock Guardrails are specifically designed to filter harmful content from chatbot inputs and responses, ensuring safe interactions for users.
* Option B: Amazon Bedrock AgentsAmazon Bedrock Agents are used to automate tasks and integrate with external tools, not to filter harmful content. This option does not meet the requirement.
* Option C: Amazon Bedrock inference APIsAmazon Bedrock inference APIs allow users to invoke foundation models for generating responses, but they do not provide built-in content filtering mechanisms.
* Option D: Amazon Bedrock custom modelsCustom models on Amazon Bedrock allow users to fine- tune models, but they do not inherently include safeguards for filtering harmful content unless explicitly implemented.
References:
AWS Bedrock User Guide: Guardrails for Responsible AI (https://docs.aws.amazon.com/bedrock/latest /userguide/guardrails.html)
AWS AI Practitioner Learning Path: Module on Responsible AI and Model Safety Amazon Bedrock Developer Guide: Building Safe AI Applications (https://aws.amazon.com/bedrock/)

NEW QUESTION: 15
A company wants to improve the accuracy of the responses from a generative AI application. The application uses a foundation model (FM) on Amazon Bedrock.
Which solution meets these requirements MOST cost-effectively?
A. Fine-tune the FM.
B. Retrain the FM.
C. Train a new FM.
D. Use prompt engineering.
Answer: D (LEAVE A REPLY)
The company wants to improve the accuracy of a generative AI application using a foundation model (FM) on Amazon Bedrock in the most cost-effective way. Prompt engineering involves optimizing the input prompts to guide the FM to produce more accurate responses without modifying the model itself. This approach is cost- effective because it does not require additional computational resources or training, unlike fine-tuning or retraining.
Exact Extract from AWS AI Documents:
From the AWS Bedrock User Guide:

"Prompt engineering is a cost-effective technique to improve the performance of foundation models. By crafting precise and context-rich prompts, users can guide the model to generate more accurate and relevant responses without the need for fine-tuning or retraining." (Source: AWS Bedrock User Guide, Prompt Engineering for Foundation Models) Detailed Explanation:
* Option A: Fine-tune the FM.Fine-tuning involves retraining the FM on a custom dataset, which requirescomputational resources, time, and cost (e.g., for Amazon Bedrock fine-tuning jobs). It is not the most cost-effective solution.
* Option B: Retrain the FM.Retraining an FM from scratch is highly resource-intensive and expensive, as it requires large datasets and significant compute power. This is not cost-effective.
* Option C: Train a new FM.Training a new FM is the most expensive option, as it involves building a model from the ground up, requiring extensive data, compute resources, and expertise. This is not cost- effective.
* Option D: Use prompt engineering.This is the correct answer. Prompt engineering adjusts the input prompts to improve the FM's responses without incurring additional compute costs, making it the most cost-effective solution for improving accuracy on Amazon Bedrock.
References:
AWS Bedrock User Guide: Prompt Engineering for Foundation Models (https://docs.aws.amazon.com /bedrock/latest/userguide/prompt-engineering.html)
AWS AI Practitioner Learning Path: Module on Generative AI Optimization Amazon Bedrock Developer Guide: Cost Optimization for Generative AI (https://aws.amazon.com/bedrock/)

NEW QUESTION: 16
A company is building a customer service chatbot. The company wants the chatbot to improve its responses by learning from past interactions and online resources.
Which AI learning strategy provides this self-improvement capability?
A. Supervised learning with a manually curated dataset of good responses and bad responses
B. Supervised learning with a continuously updated FAQ database
C. Reinforcement learning with rewards for positive customer feedback
D. Unsupervised learning to find clusters of similar customer inquiries
Answer: (SHOW ANSWER)

NEW QUESTION: 17
Which AWS service or feature can help an AI development team quickly deploy and consume a foundation model (FM) within the team's VPC?
A. Amazon Personalize
B. Amazon SageMaker JumpStart
C. PartyRock, an Amazon Bedrock Playground
D. Amazon SageMaker endpoints
Answer: B (LEAVE A REPLY)
Amazon SageMaker JumpStart is the correct service for quickly deploying and consuming a foundation model (FM) within a team's VPC.
* Amazon SageMaker JumpStart:
* Provides access to a wide range of pre-trained models and solutions that can be easily deployed and consumed within a VPC.
* Designed to simplify and accelerate the deployment of machine learning models, including foundation models.
* Why Option B is Correct:
* Rapid Deployment: JumpStart allows for quick deployment of models with minimal configuration, directly within a secure VPC environment.
* Ease of Use: Provides a user-friendly interface to select and deploy models, reducing the time to value.
* Why Other Options are Incorrect:
* A. Amazon Personalize: Focuses on creating personalized recommendations, not deploying foundation models.
* C. PartyRock: Not a recognized AWS service.
* D. Amazon SageMaker endpoints: Endpoints are for deploying specific models, not a feature for quickly starting with pre-trained foundation models.

NEW QUESTION: 18
An accounting firm wants to implement a large language model (LLM) to automate document processing.
The firm must proceed responsibly to avoid potential harms.
What should the firm do when developing and deploying the LLM? (Select TWO.)
A. Include fairness metrics for model evaluation.
B. Adjust the temperature parameter of the model.
C. Modify the training data to mitigate bias.
D. Avoid overfitting on the training data.

E. Apply prompt engineering techniques.

Answer: A,C (LEAVE A REPLY)

To implement a large language model (LLM) responsibly, the firm should focus on fairness and mitigating bias, which are critical for ethical AI deployment.

* A. Include Fairness Metrics for Model Evaluation:
* Fairness metrics help ensure that the model's predictions are unbiased and do not unfairly discriminate against any group.
* These metrics can measure disparities in model outcomes across different demographic groups, ensuring responsible AI practices.
* C. Modify the Training Data to Mitigate Bias:
* Adjusting training data to be more representative and balanced can help reduce bias in the model's predictions.
* Mitigating bias at the data level ensures that the model learns from a diverse and fair dataset, reducing potential harms in deployment.
* Why Other Options are Incorrect:
* B. Adjust the temperature parameter of the model: Controls randomness in outputs but does not directly address fairness or bias.
* D. Avoid overfitting on the training data: Important for model generalization but not directly related to responsible AI practices regarding fairness and bias.
* E. Apply prompt engineering techniques: Useful for improving model outputs but not specifically for mitigating bias or ensuring fairness.

NEW QUESTION: 19

A company is using an Amazon Bedrock base model to summarize documents for an internal use case. The company trained a custom model to improve the summarization quality.

Which action must the company take to use the custom model through Amazon Bedrock?

A. Purchase Provisioned Throughput for the custom model.

B. Deploy the custom model in an Amazon SageMaker endpoint for real-time inference.

C. Register the model with the Amazon SageMaker Model Registry.

D. Grant access to the custom model in Amazon Bedrock.

Answer: B (LEAVE A REPLY)

To use a custom model that has been trained to improve summarization quality, the company must deploy the model on an Amazon SageMaker endpoint. This allows the model to be used for real-time inference through Amazon Bedrock or other AWS services. By deploying the model in SageMaker, the custom model can be accessed programmatically via API calls, enabling integration with Amazon Bedrock.

* Option B (Correct): "Deploy the custom model in an Amazon SageMaker endpoint for real-time inference": This is the correct answer because deploying the model on SageMaker enables it to serve real-time predictions and be integrated with Amazon Bedrock.

* Option A: "Purchase Provisioned Throughput for the custom model" is incorrect because provisioned throughput is related to database or storage services, not model deployment.
* Option C: "Register the model with the Amazon SageMaker Model Registry" is incorrect because while the model registry helps with model management, it does not make the model accessible for real- time inference.
* Option D: "Grant access to the custom model in Amazon Bedrock" is incorrect because Bedrock does not directly manage custom model access; it relies on deployed endpoints like those in SageMaker.
AWS AI Practitioner References:
* Amazon SageMaker Endpoints: AWS recommends deploying models to SageMaker endpoints to use them for real-time inference in various applications.

NEW QUESTION: 20
A company is using the Generative AI Security Scoping Matrix to assess security responsibilities for its solutions. The company has identified four different solution scopes based on the matrix.
Which solution scope gives the company the MOST ownership of security responsibilities?
A. Building and training a generative AI model from scratch by using specific data that a customer owns.
B. Refining an existing third-party generative AI foundation model (FM) by fine-tuning the model by using data specific to the business.
C. Using a third-party enterprise application that has embedded generative AI features.
D. Building an application by using an existing third-party generative AI foundation model (FM).
Answer: A (LEAVE A REPLY)

NEW QUESTION: 21
Which strategy evaluates the accuracy of a foundation model (FM) that is used in image classification tasks?
A. Calculate the total cost of resources used by the model.
B. Measure the model's accuracy against a predefined benchmark dataset.
C. Count the number of layers in the neural network.
D. Assess the color accuracy of images processed by the model.
Answer: B (LEAVE A REPLY)
Measuring the model's accuracy against a predefined benchmark dataset is the correct strategy to evaluate the accuracy of a foundation model (FM) used in image classification tasks.
* Model Accuracy Evaluation:
* In image classification, the accuracy of a model is typically evaluated by comparing the predicted labels with the true labels in a benchmark dataset that is representative of the real-world data the model will encounter.

* This approach provides a quantifiable measure of how well the model performs on known data and is a standard practice in machine learning.
* Why Option B is Correct:
* Benchmarking Accuracy: Using a predefined dataset allows for consistent and reliable evaluation of model performance.
* Standard Practice: It is a widely accepted method for assessing the effectiveness of image classification models.
* Why Other Options are Incorrect:
* A. Total cost of resources: Does not measure model accuracy but rather the cost of operation.
* C. Number of layers in the neural network: Does not directly correlate with the accuracy or performance of the model.
* D. Color accuracy of images processed by the model: Is unrelated to the model's classification accuracy.

NEW QUESTION: 22
Which option describes embeddings in the context of AI?
A. A method for compressing large datasets
B. An encryption method for securing sensitive data
C. A method for visualizing high-dimensional data
D. A numerical method for data representation in a reduced dimensionality space
Answer: D (LEAVE A REPLY)
Embeddings in AI refer to numerical representations of data (e.g., text, images) in a lower-dimensional space, capturing semantic or contextual relationships. They are widely used in NLP and other AI tasks to represent complex data in a format that models can process efficiently.
Exact Extract from AWS AI Documents:
From the AWS AI Practitioner Learning Path:
"Embeddings are numerical representations of data in a reduced dimensionality space. In natural language processing, for example, word or sentence embeddings capture semantic relationships, enabling models to process text efficiently for tasks like classification or similarity search." (Source: AWS AI Practitioner Learning Path, Module on AI Concepts)
Detailed Explanation:
Option A: A method for compressing large datasetsWhile embeddings reduce dimensionality, their primary purpose is not data compression but rather to represent data in a way that preserves meaningful relationships.
This option is incorrect.
Option B: An encryption method for securing sensitive dataEmbeddings are not related to encryption or data security. They are used for data representation, making this option incorrect.

Option C: A method for visualizing high-dimensional dataWhile embeddings can sometimes be used in visualization (e.g., t-SNE), their primary role is data representation for model processing, not visualization.
This option is misleading.
Option D: A numerical method for data representation in a reduced dimensionality spaceThis is the correct answer. Embeddings transform complex data into lower-dimensional numerical vectors, preserving semantic or contextual information for use in AI models.
References:
AWS AI Practitioner Learning Path: Module on AI Concepts
Amazon Comprehend Developer Guide: Embeddings for Text Analysis (https://docs.aws.amazon.com /comprehend/latest/dg/embeddings.html)
AWS Documentation: What are Embeddings? (https://aws.amazon.com/what-is/embeddings/)

NEW QUESTION: 23
What does an F1 score measure in the context of foundation model (FM) performance?
A. Model precision and recall
B. Model speed in generating responses
C. Financial cost of operating the model
D. Energy efficiency of the model's computations
Answer: A (LEAVE A REPLY)
The F1 score is a metric used to evaluate the performance of a classification model by considering both precision and recall. Precision measures the accuracy of positive predictions (i.e., the proportion of true positive predictions among all positive predictions made by the model), while recall measures the model's ability to identify all relevant positive instances (i.e., the proportion of true positive predictions among all actual positive instances). The F1 score is the harmonic mean of precision and recall, providing a single metric that balances both concerns. This is particularly useful when dealing with imbalanced datasets or when the cost of false positives and false negatives is significant. Options B, C, and D pertain to other aspects of model performance but are not related to the F1 score.

NEW QUESTION: 24
A medical company is customizing a foundation model (FM) for diagnostic purposes. The company needs the model to be transparent and explainable to meet regulatory requirements.
Which solution will meet these requirements?
A. Configure the security and compliance by using Amazon Inspector.
B. Generate simple metrics, reports, and examples by using Amazon SageMaker Clarify.

C. Encrypt and secure training data by using Amazon Macie.
D. Gather more data. Use Amazon Rekognition to add custom labels to the data.
Answer: (SHOW ANSWER)
Amazon SageMaker Clarify provides transparency and explainability for machine learning models by generating metrics, reports, and examples that help to understand model predictions. For a medical company that needs a foundation model to be transparent and explainable to meet regulatory requirements, SageMaker Clarify is the most suitable solution.
* Amazon SageMaker Clarify:
* It helps in identifying potential bias in the data and model, and also explains model behavior by generating feature attributions, providing insights into which features are most influential in the model's predictions.
* These capabilities are critical in medical applications where regulatory compliance often mandates transparency and explainability to ensure that decisions made by the model can be trusted and audited.
* Why Option B is Correct:
* Transparency and Explainability: SageMaker Clarify is explicitly designed to provide insights into machine learning models' decision-making processes, helping meet regulatory requirements by explaining why a model made a particular prediction.
* Compliance with Regulations: The tool is suitable for use in sensitive domains, such as healthcare, where there is a need for explainable AI.
* Why Other Options are Incorrect:
* A. Amazon Inspector: Focuses on security assessments, not on explainability or model transparency.
* C. Amazon Macie: Provides data security by identifying and protecting sensitive data, but does not help in making models explainable.
* D. Amazon Rekognition: Used for image and video analysis, not relevant to making models explainable.
Thus, B is the correct answer for meeting transparency and explainability requirements for the foundation model

NEW QUESTION: 25
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

Answer: B (LEAVE A REPLY)

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 References:

* 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.

NEW QUESTION: 26

A company is using the Generative AI Security Scoping Matrix to assess security responsibilities for its solutions. The company has identified four different solution scopes based on the matrix.

Which solution scope gives the company the MOST ownership of security responsibilities?

A. Using a third-party enterprise application that has embedded generative AI features.

B. Building an application by using an existing third-party generative AI foundation model (FM).

C. Refining an existing third-party generative AI foundation model (FM) by fine-tuning the model by using data specific to the business.

D. Building and training a generative AI model from scratch by using specific data that a customer owns.

Answer: D (LEAVE A REPLY)

Building and training a generative AI model from scratch provides the company with the most ownership and control over security responsibilities. In this scenario, the company is responsible for all aspects of the security of the data, the model, and the infrastructure.

* Option D (Correct): "Building and training a generative AI model from scratch by using specific data that a customer owns": This is the correct answer because it involves

complete ownership of the model, data, and infrastructure, giving the company the highest level of responsibility for security.
* Option A: "Using a third-party enterprise application that has embedded generative AI features" is incorrect as the company has minimal control over the security of the AI features embedded within a third-party application.
* Option B: "Building an application using an existing third-party generative AI foundation model (FM)" is incorrect because security responsibilities are shared with the third-party model provider.
* Option C: "Refining an existing third-party generative AI FM by fine-tuning the model with business- specific data" is incorrect as the foundation model and part of the security responsibilities are still managed by the third party.
AWS AI Practitioner References:
* Generative AI Security Scoping Matrix on AWS: AWS provides a security responsibility matrix that outlines varying levels of control and responsibility depending on the approach to developing and using AI models.

NEW QUESTION: 27
An education provider is building a question and answer application that uses a generative AI model to explain complex concepts. The education provider wants to automatically change the style of the model response depending on who is asking the question. The education provider will give the model the age range of the user who has asked the question.
Which solution meets these requirements with the LEAST implementation effort?
A. Fine-tune the model by using additional training data that is representative of the various age ranges that the application will support.
B. Add a role description to the prompt context that instructs the model of the age range that the response should target.
C. Use chain-of-thought reasoning to deduce the correct style and complexity for a response suitable for that user.
D. Summarize the response text depending on the age of the user so that younger users receive shorter responses.
Answer: B (LEAVE A REPLY)
Adding a role description to the prompt context is a straightforward way to instruct the generative AI model to adjust its response style based on the user's age range. This method requires minimal implementation effort as it does not involve additional training or complex logic.
* Option B (Correct): "Add a role description to the prompt context that instructs the model of the age range that the response should target": This is the correct answer because it involves the least implementation effort while effectively guiding the model to tailor responses according to the age range.

* Option A: "Fine-tune the model by using additional training data" is incorrect because it requires significant effort in gathering data and retraining the model.
* Option C: "Use chain-of-thought reasoning" is incorrect as it involves complex reasoning that may not directly address the need to adjust response style based on age.
* Option D: "Summarize the response text depending on the age of the user" is incorrect because it involves additional processing steps after generating the initial response, increasing complexity.
AWS AI Practitioner References:
* Prompt Engineering Techniques on AWS: AWS recommends using prompt context effectively to guide generative models in providing tailored responses based on specific user attributes.

NEW QUESTION: 28
What does an F1 score measure in the context of foundation model (FM) performance?
A. Model precision and recall
B. Model speed in generating responses
C. Financial cost of operating the model
D. Energy efficiency of the model's computations
Answer: A (LEAVE A REPLY)
The F1 score is a metric used to evaluate the performance of a classification model by considering both precision and recall. Precision measures the accuracy of positive predictions (i.e., the proportion of true positive predictions among all positive predictions made by the model), while recall measures the model's ability to identify all relevant positive instances (i.e., the proportion of true positive predictions among all actual positive instances). The F1 score is the harmonic mean of precision and recall, providing a single metric that balances both concerns. This is particularly useful when dealing with imbalanced datasets or when the cost of false positives and false negatives is significant. Options B, C, and D pertain to other aspects of model performance but are not related to the F1 score.
Reference: AWS Certified AI Practitioner Exam Guide

NEW QUESTION: 29
A company has thousands of customer support interactions per day and wants to analyze these interactions to identify frequently asked questions and develop insights.
Which AWS service can the company use to meet this requirement?
A. Amazon Comprehend
B. Amazon Lex
C. Amazon Transcribe
D. Amazon Translate
Answer: A (LEAVE A REPLY)

NEW QUESTION: 30
An AI practitioner is using a large language model (LLM) to create content for marketing campaigns. The generated content sounds plausible and factual but is incorrect.
Which problem is the LLM having?
A. Data leakage
B. Hallucination
C. Overfitting
D. Underfitting
Answer: B (LEAVE A REPLY)
In the context of AI, "hallucination" refers to the phenomenon where a model generates outputs that are plausible-sounding but are not grounded in reality or the training data. This problem often occurs with large language models (LLMs) when they create information that sounds correct but is actually incorrect or fabricated.
* Option B (Correct): "Hallucination": This is the correct answer because the problem described involves generating content that sounds factual but is incorrect, which is characteristic of hallucination in generative AI models.
* Option A: "Data leakage" is incorrect as it involves the model accidentally learning from data it shouldn't have access to, which does not match the problem of generating incorrect content.
* Option C: "Overfitting" is incorrect because overfitting refers to a model that has learned the training data too well, including noise, and performs poorly on new data.
* Option D: "Underfitting" is incorrect because underfitting occurs when a model is too simple to capture the underlying patterns in the data, which is not the issue here.
AWS AI Practitioner References:
* Large Language Models on AWS: AWS discusses the challenge of hallucination in large language models and emphasizes techniques to mitigate it, such as using guardrails and fine-tuning.

NEW QUESTION: 31
A company wants to use AI to protect its application from threats. The AI solution needs to check if an IP address is from a suspicious source.
Which solution meets these requirements?
A. Build a speech recognition system.
B. Create a natural language processing (NLP) named entity recognition system.
C. Develop an anomaly detection system.
D. Create a fraud forecasting system.
Answer: C (LEAVE A REPLY)
An anomaly detection system is suitable for identifying unusual patterns or behaviors, such as suspicious IP addresses, which might indicate a potential threat.
* Anomaly Detection:

* Anomaly detection uses machine learning algorithms to identify deviations from normal behavior, such as unexpected traffic from a suspicious IP address.
* This is a common approach for identifying potential threats or malicious activity in cybersecurity applications.
* Why Option C is Correct:
* Detects Suspicious Behavior: An anomaly detection system can monitor and detect IP addresses that exhibit unusual or suspicious patterns.
* Real-time Monitoring: Provides continuous analysis of network traffic to identify potential security threats.
* Why Other Options are Incorrect:
* A. Speech recognition system: Is unrelated to detecting suspicious IP addresses.
* B. NLP named entity recognition: Focuses on identifying entities in text, not IP address analysis.
* D. Fraud forecasting system: Generally used for predicting fraud, not directly applicable to identifying suspicious IPs.
Thus, C is the correct answer for detecting suspicious IP addresses.

NEW QUESTION: 32
A company has terabytes of data in a database that the company can use for business analysis. The company wants to build an AI-based application that can build a SQL query from input text that employees provide. The employees have minimal experience with technology.
Which solution meets these requirements?
A. Support vector machine
B. WaveNet
C. Generative pre-trained transformers (GPT)
D. Residual neural network
Answer: C (LEAVE A REPLY)

NEW QUESTION: 33

A company wants to deploy a conversational chatbot to answer customer questions. The chatbot is based on a fine-tuned Amazon SageMaker JumpStart model. The application must comply with multiple regulatory frameworks.
Which capabilities can the company show compliance for? (Select TWO.)
A. Auto scaling inference endpoints
B. Threat detection
C. Data protection
D. Cost optimization
E. Loosely coupled microservices
Answer: B,C (LEAVE A REPLY)
To comply with multiple regulatory frameworks, the company must ensure data protection and threat detection. Data protection involves safeguarding sensitive customer information, while threat detection identifies and mitigates security threats to the application.
* Option C (Correct): "Data protection": This is correct because data protection is critical for compliance with privacy and security regulations.
* Option B (Correct): "Threat detection": This is correct because detecting and mitigating threats is essential to maintaining the security posture required for regulatory compliance.
* Option A: "Auto scaling inference endpoints" is incorrect because auto-scaling does not directly relate to regulatory compliance.
* Option D: "Cost optimization" is incorrect because it is focused on managing expenses, not compliance.
* Option E: "Loosely coupled microservices" is incorrect because this architectural approach does not directly address compliance requirements.
AWS AI Practitioner References:
* AWS Compliance Capabilities: AWS offers services and tools, such as data protection and threat detection, to help companies meet regulatory requirements for security and privacy.

NEW QUESTION: 34
A financial institution is using Amazon Bedrock to develop an AI application. The application is hosted in a VPC. To meet regulatory compliance standards, the VPC is not allowed access to any internet traffic.
Which AWS service or feature will meet these requirements?
A. Amazon Macie
B. Internet gateway
C. AWS PrivateLink
D. Amazon CloudFront
Answer: C (LEAVE A REPLY)

NEW QUESTION: 35

A company is implementing the Amazon Titan foundation model (FM) by using Amazon Bedrock. The company needs to supplement the model by using relevant data from the company's private data sources.
Which solution will meet this requirement?
A. Use a different FM
B. Choose a lower temperature value
C. Create an Amazon Bedrock knowledge base
D. Enable model invocation logging
Answer: C (LEAVE A REPLY)
Creating an Amazon Bedrock knowledge base allows the integration of external or private data sources with a foundation model (FM) like Amazon Titan. This integration helps supplement the model with relevant data from the company's private data sources to enhance its responses.
* Option C (Correct): "Create an Amazon Bedrock knowledge base": This is the correct answer as it enables the company to incorporate private data into the FM to improve its effectiveness.
* Option A: "Use a different FM" is incorrect because it does not address the need to supplement the current model with private data.
* Option B: "Choose a lower temperature value" is incorrect as it affects output randomness, not the integration of private data.
* Option D: "Enable model invocation logging" is incorrect because logging does not help in supplementing the model with additional data.
AWS AI Practitioner References:
* Amazon Bedrock and Knowledge Integration: AWS explains how creating a knowledge base allows Amazon Bedrock to use external data sources to improve the FM's relevance and accuracy.

NEW QUESTION: 36
A large retailer receives thousands of customer support inquiries about products every day. The customer support inquiries need to be processed and responded to quickly. The company wants to implement Agents for Amazon Bedrock.
What are the key benefits of using Amazon Bedrock agents that could help this retailer?
A. Generation of custom foundation models (FMs) to predict customer needs
B. Automation of repetitive tasks and orchestration of complex workflows
C. Automatically calling multiple foundation models (FMs) and consolidating the results
D. Selecting the foundation model (FM) based on predefined criteria and metrics
Answer: B (LEAVE A REPLY)
Amazon Bedrock Agents provide the capability to automate repetitive tasks and orchestrate complex workflows using generative AI models. This is particularly beneficial for customer support inquiries, where quick and efficient processing is crucial.

* Option B (Correct): "Automation of repetitive tasks and orchestration of complex workflows":
This is the correct answer because Bedrock Agents can automate common customer service tasks and streamline complex processes, improving response times and efficiency.
* Option A: "Generation of custom foundation models (FMs) to predict customer needs" is incorrect as Bedrock agents do not create custom models.
* Option C: "Automatically calling multiple foundation models (FMs) and consolidating the results" is incorrect because Bedrock agents focus on task automation rather than combining model outputs.
* Option D: "Selecting the foundation model (FM) based on predefined criteria and metrics" is incorrect as Bedrock agents are not designed for selecting models.
AWS AI Practitioner References:
* Amazon Bedrock Documentation: AWS explains that Bedrock Agents automate tasks and manage complex workflows, making them ideal for customer support automation.

NEW QUESTION: 37
In which stage of the generative AI model lifecycle are tests performed to examine the model's accuracy?
A. Deployment
B. Data selection
C. Fine-tuning
D. Evaluation
Answer: (SHOW ANSWER)
The evaluation stage of the generative AI model lifecycle involves testing the model to assess its performance, including accuracy, coherence, and other metrics. This stage ensures the model meets the desired quality standards before deployment.
Exact Extract from AWS AI Documents:
From the AWS AI Practitioner Learning Path:
"The evaluation phase in the machine learning lifecycle involves testing the model against validation or test datasets to measure its performance metrics, such as accuracy, precision, recall, or task-specific metrics for generative AI models." (Source: AWS AI Practitioner Learning Path, Module on Machine Learning Lifecycle) Detailed Explanation:
* Option A: DeploymentDeployment involves making the model available for use in production. While monitoring occurs post-deployment, accuracy testing is performed earlier in the evaluation stage.
* Option B: Data selectionData selection involves choosing and preparing data for training, not testing the model's accuracy.
* Option C: Fine-tuningFine-tuning adjusts a pre-trained model to improve performance for a specific task, but it is not the stage where accuracy is formally tested.

* Option D: EvaluationThis is the correct answer. The evaluation stage is where tests are conducted to examine the model's accuracy and other performance metrics, ensuring it meets requirements.
References:
AWS AI Practitioner Learning Path: Module on Machine Learning Lifecycle Amazon SageMaker Developer Guide: Model Evaluation (https://docs.aws.amazon.com/sagemaker/latest/dg/model-evaluation.html)
AWS Documentation: Generative AI Lifecycle (https://aws.amazon.com/machine-learning/)

NEW QUESTION: 38
A company wants to make a chatbot to help customers. The chatbot will help solve technical problems without human intervention. The company chose a foundation model (FM) for the chatbot. The chatbot needs to produce responses that adhere to company tone.
Which solution meets these requirements?
A. Define a higher number for the temperature parameter.
B. Experiment and refine the prompt until the FM produces the desired responses.
C. Use batch inferencing to process detailed responses.
D. Set a low limit on the number of tokens the FM can produce.
Answer: (SHOW ANSWER)

NEW QUESTION: 39
A digital devices company wants to predict customer demand for memory hardware. The company does not have coding experience or knowledge of ML algorithms and needs to develop a data-driven predictive model.
The company needs to perform analysis on internal data and external data.
Which solution will meet these requirements?
A. Store the data in Amazon S3. Create ML models and demand forecast predictions by using Amazon SageMaker built-in algorithms that use the data from Amazon S3.
B. Import the data into Amazon SageMaker Data Wrangler. Create ML models and demand forecast predictions by using SageMaker built-in algorithms.
C. Import the data into Amazon SageMaker Data Wrangler. Build ML models and demand forecast predictions by using an Amazon Personalize Trending-Now recipe.
D. Import the data into Amazon SageMaker Canvas. Build ML models and demand forecast predictions by selecting the values in the data from SageMaker Canvas.
Answer: D (LEAVE A REPLY)
Amazon SageMaker Canvas is a visual, no-code machine learning interface that allows users to build machine learning models without having any coding experience or knowledge of machine learning algorithms. It enables users to analyze internal and external data, and make predictions using a guided interface.

* Option D (Correct): "Import the data into Amazon SageMaker Canvas. Build ML models and demand forecast predictions by selecting the values in the data from SageMaker Canvas": This is the correct answer because SageMaker Canvas is designed for users without coding experience, providing a visual interface to build predictive models with ease.
* Option A: "Store the data in Amazon S3 and use SageMaker built-in algorithms" is incorrect because it requires coding knowledge to interact with SageMaker's built-in algorithms.
* Option B: "Import the data into Amazon SageMaker Data Wrangler" is incorrect. Data Wrangler is primarily for data preparation and not directly focused on creating ML models without coding.
* Option C: "Use Amazon Personalize Trending-Now recipe" is incorrect as Amazon Personalize is for building recommendation systems, not for general demand forecasting.
AWS AI Practitioner References:
* Amazon SageMaker Canvas Overview: AWS documentation emphasizes Canvas as a no-code solution for building machine learning models, suitable for business analysts and users with no coding experience.

NEW QUESTION: 40
A company has developed an ML model for image classification. The company wants to deploy the model to production so that a web application can use the model.
The company needs to implement a solution to host the model and serve predictions without managing any of the underlying infrastructure.
Which solution will meet these requirements?
A. Use AWS Batch to host the model and serve predictions.
B. Use Amazon SageMaker Serverless Inference to deploy the model.
C. Use Amazon CloudFront to deploy the model.
D. Use Amazon API Gateway to host the model and serve predictions.
Answer: B (LEAVE A REPLY)

NEW QUESTION: 41
A company needs to train an ML model to classify images of different types of animals. The company has a large dataset of labeled images and will not label more data. Which type of learning should the company use to train the model?
A. Supervised learning.
B. Unsupervised learning.
C. Reinforcement learning.
D. Active learning.
Answer: A (LEAVE A REPLY)
Supervised learning is appropriate when the dataset is labeled. The model uses this data to learn patterns and classify images. Unsupervised learning, reinforcement learning, and

active learning are not suitable since they either require unlabeled data or different problem settings. References: AWS Machine Learning Best Practices.

NEW QUESTION: 42
A company is building a contact center application and wants to gain insights from customer conversations. The company wants to analyze and extract key information from the audio of the customer calls.
Which solution meets these requirements?
A. Create classification labels by using Amazon Comprehend.
B. Transcribe call recordings by using Amazon Transcribe.
C. Extract information from call recordings by using Amazon SageMaker Model Monitor.
D. Build a conversational chatbot by using Amazon Lex.
Answer: B (LEAVE A REPLY)

NEW QUESTION: 43
A company's large language model (LLM) is experiencing hallucinations.
How can the company decrease hallucinations?
A. Set up Agents for Amazon Bedrock to supervise the model training.
B. Use data pre-processing and remove any data that causes hallucinations.
C. Decrease the temperature inference parameter for the model.
D. Use a foundation model (FM) that is trained to not hallucinate.
Answer: C (LEAVE A REPLY)
Hallucinations in large language models (LLMs) occur when the model generates outputs that are factually incorrect, irrelevant, or not grounded in the input data. To mitigate hallucinations, adjusting the model's inference parameters, particularly the temperature, is a well-documented approach in AWS AI Practitioner resources. The temperature parameter controls the randomness of the model's output. A lower temperature makes the model more deterministic, reducing the likelihood of generating creative but incorrect responses, which are often the cause of hallucinations.
Exact Extract from AWS AI Documents:
From the AWS documentation on Amazon Bedrock and LLMs:
"The temperature parameter controls the randomness of the generated text. Higher values (e.g., 0.8 or above) increase creativity but may lead to less coherent or factually incorrect outputs, while lower values (e.g., 0.2 or 0.3) make the output more focused and deterministic, reducing the likelihood of hallucinations." (Source: AWS Bedrock User Guide, Inference Parameters for Text Generation) Detailed Explanation:
* Option A: Set up Agents for Amazon Bedrock to supervise the model training.Agents for Amazon Bedrock are used to automate tasks and integrate LLMs with external tools, not to supervise model training or directly address hallucinations. This option is incorrect as it does not align with the purpose of Agents in Bedrock.

* Option B: Use data pre-processing and remove any data that causes hallucinations.While data pre- processing can improve model performance, identifying and removing specific data that causes hallucinations is impractical because hallucinations are often a result of the model's generative process rather than specific problematic data points. This approach is not directly supported by AWS documentation for addressing hallucinations.
* Option C: Decrease the temperature inference parameter for the model.This is the correct approach. Lowering the temperature reduces the randomness in the model's output, making it more likely to stick to factual and contextually relevant responses. AWS documentation explicitly mentions adjusting inference parameters like temperature to control output quality and mitigate issues like hallucinations.
* Option D: Use a foundation model (FM) that is trained to not hallucinate.No foundation model is explicitly trained to "not hallucinate," as hallucinations are an inherent challenge in LLMs. While some models may be fine-tuned for specific tasks to reduce hallucinations, this is not a standard feature of foundation models available on Amazon Bedrock.
References:
AWS Bedrock User Guide: Inference Parameters for Text Generation (https://docs.aws.amazon.com/bedrock /latest/userguide/model-parameters.html)
AWS AI Practitioner Learning Path: Module on Large Language Models and Inference Configuration Amazon Bedrock Developer Guide: Managing Model Outputs (https://docs.aws.amazon.com/bedrock/latest /devguide/)

NEW QUESTION: 44
Which AW5 service makes foundation models (FMs) available to help users build and scale generative AI applications?
A. Amazon Q Developer
B. Amazon Bedrock
C. Amazon Kendra
D. Amazon Comprehend
Answer: B (LEAVE A REPLY)
Amazon Bedrock is a fully managed service that provides access to foundation models (FMs) from various providers, enabling users to build and scale generative AI applications. It simplifies the process of integrating FMs into applications for tasks like text generation, chatbots, and more.
Exact Extract from AWS AI Documents:
From the AWS Bedrock User Guide:
"Amazon Bedrock is a fully managed service that makes foundation models (FMs) from leading AI providers available through a single API, enabling developers to build and scale generative AI applications with ease." (Source: AWS Bedrock User Guide, Introduction to Amazon Bedrock) Detailed Explanation:

* Option A: Amazon Q DeveloperAmazon Q Developer is an AI-powered assistant for coding and AWS service guidance, not a service for hosting or providing foundation models.
* Option B: Amazon BedrockThis is the correct answer. Amazon Bedrock provides access to foundation models, making it the primary service for building and scaling generative AI applications.
* Option C: Amazon KendraAmazon Kendra is an intelligent search service powered by machine learning, not a service for providing foundation models or building generative AI applications.
* Option D: Amazon ComprehendAmazon Comprehend is an NLP service for text analysis tasks like sentiment analysis, not for providing foundation models or supporting generative AI.
References:
AWS Bedrock User Guide: Introduction to Amazon Bedrock (https://docs.aws.amazon.com/bedrock/latest /userguide/what-is-bedrock.html)
AWS AI Practitioner Learning Path: Module on Generative AI Services
AWS Documentation: Generative AI on AWS (https://aws.amazon.com/generative-ai/)

NEW QUESTION: 45
Which AWS feature records details about ML instance data for governance and reporting?
A. Amazon SageMaker Model Cards
B. Amazon SageMaker Debugger
C. Amazon SageMaker Model Monitor
D. Amazon SageMaker JumpStart
Answer: (SHOW ANSWER)
Amazon SageMaker Model Cards provide a centralized and standardized repository for documenting machine learning models. They capture key details such as the model's intended use, training and evaluation datasets, performance metrics, ethical considerations, and other relevant information. This documentation facilitates governance and reporting by ensuring that all stakeholders have access to consistent and comprehensive information about each model. While Amazon SageMaker Debugger is used for real-time debugging and monitoring during training, and Amazon SageMaker Model Monitor tracks deployed models for data and prediction quality, neither offers the comprehensive documentation capabilities of Model Cards. Amazon SageMaker JumpStart provides pre-built models and solutions but does not focus on governance documentation.

NEW QUESTION: 46
An e-commerce company wants to build a solution to determine customer sentiments based on written customer reviews of products.

Which AWS services meet these requirements? (Select TWO.)
A. Amazon Lex
B. Amazon Comprehend
C. Amazon Polly
D. Amazon Bedrock
E. Amazon Rekognition
Answer: B,D (LEAVE A REPLY)
To determine customer sentiments based on written customer reviews, the company can use Amazon Comprehend and Amazon Bedrock.
* Amazon Comprehend:
* A natural language processing (NLP) service that uses machine learning to uncover insights and relationships in text.
* Can analyze customer reviews to detect sentiments (positive, negative, neutral, or mixed) automatically.
* Amazon Bedrock:
* Provides access to foundational models (FMs) from multiple AI companies for tasks such as text generation, summarization, and sentiment analysis.
* The company can use a pre-trained sentiment analysis model available on Amazon Bedrock for processing customer reviews.
* Why Other Options are Incorrect:
* A. Amazon Lex: Used for building conversational interfaces like chatbots, not for sentiment analysis.
* C. Amazon Polly: Converts text to speech; it doesn't analyze sentiment.
* E. Amazon Rekognition: Analyzes images and videos, not text.

NEW QUESTION: 47
An AI practitioner trained a custom model on Amazon Bedrock by using a training dataset that contains confidential dat a. 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. Mask the confidential data in the inference responses by using dynamic data masking.
B. Encrypt the confidential data in the inference responses by using Amazon SageMaker.

C. Delete the custom model. Remove the confidential data from the training dataset. Retrain the custom model.
D. Encrypt the confidential data in the custom model by using AWS Key Management Service (AWS KMS).
Answer: (SHOW ANSWER)

NEW QUESTION: 48
A company wants to create a new solution by using AWS Glue. The company has minimal programming experience with AWS Glue.
Which AWS service can help the company use AWS Glue?
A. Amazon Q Developer
B. AWS Config
C. Amazon Personalize
D. Amazon Comprehend
Answer: A (LEAVE A REPLY)
AWS Glue is a serverless data integration service that enables users to extract, transform, and load (ETL) data. For a company with minimal programming experience, Amazon Q Developer provides an AI-powered assistant that can generate code, explain AWS services, and guide users through tasks like creating AWS Glue jobs. This makes it an ideal tool to help the company use AWS Glue effectively.
Exact Extract from AWS AI Documents:
From the AWS Documentation on Amazon Q Developer:
"Amazon Q Developer is an AI-powered assistant that helps developers by generating code, answering questions about AWS services, and providing step-by-step guidance for tasks such as building ETL pipelines with AWS Glue. It is designed to assist users with varying levels of expertise, including those with minimal programming experience." (Source: AWS Documentation, Amazon Q Developer Overview) Detailed Explanation:
Option A: Amazon Q Developer
This is the correct answer. Amazon Q Developer can assist the company by generating AWS Glue scripts, explaining Glue concepts, and providing guidance on setting up ETL jobs, which is particularly helpful for users with limited programming experience.
Option B: AWS Config
AWS Config is used for tracking and managing resource configurations and compliance, not for assisting with coding or using services like AWS Glue. This option is incorrect.
Option C: Amazon Personalize
Amazon Personalize is a machine learning service for building recommendation systems, not for assisting with data integration or AWS Glue. This option is irrelevant.
Option D: Amazon Comprehend
Amazon Comprehend is an NLP service for analyzing text, not for helping users write code or use AWS Glue.

This option does not meet the requirements.
References:
AWS Documentation: Amazon Q Developer Overview (https://aws.amazon.com/q/developer/) AWS Glue Developer Guide: Introduction to AWS Glue (https://docs.aws.amazon.com/glue/latest/dg/what-is- glue.html) AWS AI Practitioner Learning Path: Module on AWS Developer Tools and Services

NEW QUESTION: 49
An AI practitioner wants to use a foundation model (FM) to design a search application. The search application must handle queries that have text and images.
Which type of FM should the AI practitioner use to power the search application?
A. Multi-modal embedding model
B. Text embedding model
C. Multi-modal generation model
D. Image generation model
Answer: A (LEAVE A REPLY)
A multi-modal embedding model is the correct type of foundation model (FM) for powering a search application that handles queries containing both text and images.
* Multi-Modal Embedding Model:
* Can process and integrate different types of data (e.g., text and images) into a common representation space, enabling a unified search capability.
* Suitable for applications where queries or content involve multiple data modalities.
* Why Option A is Correct:
* Handles Multiple Modalities: Supports both text and image data, aligning with the application's requirement.
* Improves Search Relevance: Allows for more accurate and relevant search results across different types of input data.
* Why Other Options are Incorrect:
* B. Text embedding model: Only handles text data, not images.
* C. Multi-modal generation model: Focuses on generating outputs rather than embedding for search tasks.
* D. Image generation model: Only handles image data, not suitable for text queries.

NEW QUESTION: 50
A company is developing a new model to predict the prices of specific items. The model performed well on the training dataset. When the company deployed the model to production, the model's performance decreased significantly.
What should the company do to mitigate this problem?
A. Reduce the volume of data that is used in training.
B. Add hyperparameters to the model.
C. Increase the volume of data that is used in training.

D. Increase the model training time.

Answer: (SHOW ANSWER)

When a model performs well on the training data but poorly in production, it is often due to overfitting.

Overfitting occurs when a model learns patterns and noise specific to the training data, which does not generalize well to new, unseen data in production. Increasing the volume of data used in training can help mitigate this problem by providing a more diverse and representative dataset, which helps the model generalize better.

* Option C (Correct): "Increase the volume of data that is used in training": Increasing the data volume can help the model learn more generalized patterns rather than specific features of the training dataset, reducing overfitting and improving performance in production.
* Option A: "Reduce the volume of data that is used in training" is incorrect, as reducing data volume would likely worsen the overfitting problem.
* Option B: "Add hyperparameters to the model" is incorrect because adding hyperparameters alone does not address the issue of data diversity or model generalization.
* Option D: "Increase the model training time" is incorrect because simply increasing training time does not prevent overfitting; the model needs more diverse data.

AWS AI Practitioner References:

* Best Practices for Model Training on AWS: AWS recommends using a larger and more diverse training dataset to improve a model's generalization capability and reduce the risk of overfitting.

NEW QUESTION: 51

A company has documents that are missing some words because of a database error. The company wants to build an ML model that can suggest potential words to fill in the missing text.

Which type of model meets this requirement?

A. Clustering models

B. Topic modeling

C. BERT-based models

D. Prescriptive ML models

Answer: C (LEAVE A REPLY)

NEW QUESTION: 52

A company wants to build a lead prioritization application for its employees to contact potential customers.

The application must give employees the ability to view and adjust the weights assigned to different variables in the model based on domain knowledge and expertise.

Which ML model type meets these requirements?

A. Logistic regression model
B. Deep learning model built on principal components
C. K-nearest neighbors (k-NN) model
D. Neural network
Answer: A (LEAVE A REPLY)
The company needs an ML model for a lead prioritization application where employees can view and adjust the weights assigned to different variables based on domain knowledge. Logistic regression is a linear model that assigns interpretable weights to input features, making it easy for users to understand and modify these weights. This interpretability and adjustability make it suitable for the requirements.
Exact Extract from AWS AI Documents:
From the AWS AI Practitioner Learning Path:
"Logistic regression is a supervised learning algorithm used for classification tasks. It is highly interpretable, as it assigns weights to each feature, allowing users to understand and adjust the importance of different variables based on domain expertise." (Source: AWS AI Practitioner Learning Path, Module on Machine Learning Algorithms) Detailed Explanation:
* Option A: Logistic regression modelThis is the correct answer. Logistic regression provides interpretable coefficients (weights) for each feature, enabling employees to view and adjust them based on domain knowledge, meeting the application's requirements.
* Option B: Deep learning model built on principal componentsDeep learning models, even when using principal components, are complex and lack interpretability. The weights in such models are not easily adjustable by users, making this option unsuitable.
* Option C: K-nearest neighbors (k-NN) modelk-NN is a non-parametric model that does not assign explicit weights to features. It relieson distance metrics, which are not easily adjustable based on domain knowledge, so it does not meet the requirements.
* Option D: Neural networkNeural networks are highly complex and lack interpretability, as their weights are not directly tied to input features in a human-understandable way. Adjusting weights based on domain knowledge is impractical, making this option incorrect.
References:
AWS AI Practitioner Learning Path: Module on Machine Learning Algorithms Amazon SageMaker Developer Guide: Logistic Regression
(https://docs.aws.amazon.com/sagemaker/latest/dg
/algos.html)
AWS Documentation: Interpretable Machine Learning Models
(https://aws.amazon.com/machine-learning/)

NEW QUESTION: 53
An AI practitioner is using a large language model (LLM) to create content for marketing campaigns. The generated content sounds plausible and factual but is incorrect.
Which problem is the LLM having?

A. Hallucination
B. Data leakage
C. Underfitting
D. Overfitting
Answer: A (LEAVE A REPLY)

NEW QUESTION: 54
A company needs to build its own large language model (LLM) based on only the company's private data.
The company is concerned about the environmental effect of the training process.
Which Amazon EC2 instance type has the LEAST environmental effect when training LLMs?
A. Amazon EC2 C series
B. Amazon EC2 G series
C. Amazon EC2 P series
D. Amazon EC2 Trn series
Answer: (SHOW ANSWER)
The Amazon EC2 Trn series (Trainium) instances are designed for high-performance, cost-effective machine learning training while being energy-efficient. AWS Trainium-powered instances are optimized for deep learning models and have been developed to minimize environmental impact by maximizing energy efficiency.
* Option D (Correct): "Amazon EC2 Trn series": This is the correct answer because the Trn series is purpose-built for training deep learning models with lower energy consumption, which aligns with the company's concern about environmental effects.
* Option A: "Amazon EC2 C series" is incorrect because it is intended for compute-intensive tasks but not specifically optimized for ML training with environmental considerations.
* Option B: "Amazon EC2 G series" (Graphics Processing Unit instances) is optimized for graphics- intensive applications but does not focus on minimizing environmental impact for training.
* Option C: "Amazon EC2 P series" is designed for ML training but does not offer the same level of energy efficiency as the Trn series.
AWS AI Practitioner References:
* AWS Trainium Overview: AWS promotes Trainium instances as their most energy-efficient and cost- effective solution for ML model training.

NEW QUESTION: 55
An AI practitioner has a database of animal photos. The AI practitioner wants to automatically identify and categorize the animals in the photos without manual human effort.
Which strategy meets these requirements?

A. Object detection
B. Anomaly detection
C. Named entity recognition
D. Inpainting
Answer: A (LEAVE A REPLY)
Object detection is the correct strategy for automatically identifying and categorizing animals in photos.
* Object Detection:
* A computer vision technique that identifies and locates objects within an image and assigns them to predefined categories.
* Ideal for tasks such as identifying animals in photos, where the goal is to detect specific objects (animals) and categorize them accordingly.
* Why Option A is Correct:
* Automatic Identification: Object detection models can automatically identify different types of animals in the images without manual intervention.
* Categorization Capability: Assigns labels to detected objects, fulfilling the requirement for categorizing animals.
* Why Other Options are Incorrect:
* B. Anomaly detection: Identifies outliers or unusual patterns, not specific objects in images.
* C. Named entity recognition: Used in NLP to identify entities in text, not for image processing.
* D. Inpainting: Used for filling in missing parts of images, not for detecting or categorizing objects.

NEW QUESTION: 56
A company is building an ML model. The company collected new data and analyzed the data by creating a correlation matrix, calculating statistics, and visualizing the data.
Which stage of the ML pipeline is the company currently in?
A. Hyperparameter tuning
B. Feature engineering
C. Data pre-processing
D. Exploratory data analysis
Answer: (SHOW ANSWER)

NEW QUESTION: 57
A company needs to train an ML model to classify images of different types of animals. The company has a large dataset of labeled images and will not label more data. Which type of learning should the company use to train the model?
A. Active learning.
B. Reinforcement learning.

C. Unsupervised learning.
D. Supervised learning.
Answer: D (LEAVE A REPLY)
Supervised learning is appropriate when the dataset is labeled. The model uses this data to learn patterns and classify images. Unsupervised learning, reinforcement learning, and active learning are not suitable since they either require unlabeled data or different problem settings. References: AWS Machine Learning Best Practices.

NEW QUESTION: 58
An AI practitioner is building a model to generate images of humans in various professions. The AI practitioner discovered that the input data is biased and that specific attributes affect the image generation and create bias in the model.
Which technique will solve the problem?
A. Retrieval Augmented Generation (RAG)
B. Watermark detection for images
C. Data augmentation for imbalanced classes
D. Model monitoring for class distribution
Answer: (SHOW ANSWER)

NEW QUESTION: 59
A company is using Amazon SageMaker Studio notebooks to build and train ML models. The company stores the data in an Amazon S3 bucket. The company needs to manage the flow of data from Amazon S3 to SageMaker Studio notebooks.
Which solution will meet this requirement?
A. Use Amazon Inspector to monitor SageMaker Studio.
B. Configure SageMaker to use a VPC with an S3 endpoint.
C. Use Amazon Macie to monitor SageMaker Studio.
D. Configure SageMaker to use S3 Glacier Deep Archive.
Answer: B (LEAVE A REPLY)

NEW QUESTION: 60
A student at a university is copying content from generative AI to write essays.
Which challenge of responsible generative AI does this scenario represent?
A. Privacy
B. Toxicity
C. Hallucinations
D. Plagiarism
Answer: D (LEAVE A REPLY)

NEW QUESTION: 61

An airline company wants to build a conversational AI assistant to answer customer questions about flight schedules, booking, and payments. The company wants to use large language models (LLMs) and a knowledge base to create a text-based chatbot interface. Which solution will meet these requirements with the LEAST development effort?

A. Train models on Amazon SageMaker Autopilot.

B. Develop a Retrieval Augmented Generation (RAG) agent by using Amazon Bedrock.

C. Create a Python application by using Amazon Q Developer.

D. Fine-tune models on Amazon SageMaker Jumpstart.

Answer: B (LEAVE A REPLY)

The airline company aims to build a conversational AI assistant using large language models (LLMs) and a knowledge base to create a text-based chatbot with minimal development effort. Retrieval Augmented Generation (RAG) on Amazon Bedrock is an ideal solution because it combines LLMs with a knowledge base to provide accurate, contextually relevant responses without requiring extensive model training or custom development. RAG retrieves relevant information from a knowledge base and uses an LLM to generate responses, simplifying the development process.

Exact Extract from AWS AI Documents:

From the AWS Bedrock User Guide:

"Retrieval Augmented Generation (RAG) in Amazon Bedrock enables developers to build conversational AI applications by combining foundation models with external knowledge bases. This approach minimizes development effort by leveraging pre-trained models and integrating them with data sources, such as FAQs or databases, to provide accurate and contextually relevant responses." (Source: AWS Bedrock User Guide, Retrieval Augmented Generation) Detailed Explanation:

* Option A: Train models on Amazon SageMaker Autopilot.SageMaker Autopilot is designed for automated machine learning (AutoML) tasks like classification or regression, not for building conversational AI with LLMs and knowledge bases. It requires significant data preparation and is not optimized for chatbot development, making it less suitable.

* Option B: Develop a Retrieval Augmented Generation (RAG) agent by using Amazon Bedrock.

This is the correct answer. RAG on Amazon Bedrock allows the company to use pre-trained LLMs and integrate them with a knowledge base (e.g., flight schedules or FAQs) to build a chatbot with minimal effort. It avoids the need for extensive training or coding, aligning with the requirement for least development effort.

* Option C: Create a Python application by using Amazon Q Developer.While Amazon Q Developer can assist with code generation, building a chatbot from scratch in Python requires significant development effort, including integrating LLMs and a knowledge base manually, which is more complex than using RAG on Bedrock.

* Option D: Fine-tune models on Amazon SageMaker Jumpstart.Fine-tuning models on SageMaker Jumpstart requires preparing training data and customizing LLMs, which

involves more effort than using a pre-built RAG solution on Bedrock. This option is not the least effort-intensive.
References:
AWS Bedrock User Guide: Retrieval Augmented Generation (https://docs.aws.amazon.com/bedrock/latest/userguide/rag.html)
AWS AI Practitioner Learning Path: Module on Generative AI and Conversational AI
Amazon Bedrock Developer Guide: Building Conversational AI (https://aws.amazon.com/bedrock/)

NEW QUESTION: 62
A company wants to use generative AI to increase developer productivity and software development. The company wants to use Amazon Q Developer.
What can Amazon Q Developer do to help the company meet these requirements?
A. Create software snippets, reference tracking, and open-source license tracking.
B. Run an application without provisioning or managing servers.
C. Enable voice commands for coding and providing natural language search.
D. Convert audio files to text documents by using ML models.
Answer: A (LEAVE A REPLY)
Amazon Q Developer is a tool designed to assist developers in increasing productivity by generating code snippets, managing reference tracking, and handling open-source license tracking. These features help developers by automating parts of the software development process.
* Option A (Correct): "Create software snippets, reference tracking, and open-source license tracking": This is the correct answer because these are key features that help developers streamline and automate tasks, thus improving productivity.
* Option B: "Run an application without provisioning or managing servers" is incorrect as it refers to AWS Lambda or AWS Fargate, not Amazon Q Developer.
* Option C: "Enable voice commands for coding and providing natural language search" is incorrect because this is not a function of Amazon Q Developer.
* Option D: "Convert audio files to text documents by using ML models" is incorrect as this refers to Amazon Transcribe, not Amazon Q Developer.

AWS AI Practitioner References:
* Amazon Q Developer Features: AWS documentation outlines how Amazon Q Developer supports developers by offering features that reduce manual effort and improve efficiency.

NEW QUESTION: 63
A company wants to display the total sales for its top-selling products across various retail locations in the past 12 months.
Which AWS solution should the company use to automate the generation of graphs?
A. Amazon Q in Amazon QuickSight
B. Amazon Q in Amazon EC2
C. Amazon Q in AWS Chatbot
D. Amazon Q Developer
Answer: A (LEAVE A REPLY)

NEW QUESTION: 64
A company wants to classify human genes into 20 categories based on gene characteristics. The company needs an ML algorithm to document how the inner mechanism of the model affects the output.
Which ML algorithm meets these requirements?
A. Decision trees
B. Linear regression
C. Logistic regression
D. Neural networks
Answer: A (LEAVE A REPLY)
Decision trees are an interpretable machine learning algorithm that clearly documents the decision-making process by showing how each input feature affects the output. This transparency is particularly useful when explaining how the model arrives at a certain decision, making it suitable for classifying genes into categories.
* Option A (Correct): "Decision trees": This is the correct answer because decision trees provide a clear and interpretable representation of how input features influence the model's output, making it ideal for understanding the inner mechanisms affecting predictions.
* Option B: "Linear regression" is incorrect because it is used for regression tasks, not classification.
* Option C: "Logistic regression" is incorrect as it does not provide the same level of interpretability in documenting decision-making processes.
* Option D: "Neural networks" is incorrect because they are often considered "black boxes" and do not easily explain how they arrive at their outputs.
AWS AI Practitioner References:
* Interpretable Machine Learning Models on AWS: AWS supports using interpretable models, such as decision trees, for tasks that require clear documentation of how input data affects output decisions.

NEW QUESTION: 65
An e-commerce company wants to build a solution to determine customer sentiments based on written customer reviews of products.
Which AWS services meet these requirements? (Select TWO.)
A. Amazon Rekognition
B. Amazon Bedrock
C. Amazon Polly
D. Amazon Lex
E. Amazon Comprehend
Answer: B,E (LEAVE A REPLY)

NEW QUESTION: 66
An education provider is building a question and answer application that uses a generative AI model to explain complex concepts. The education provider wants to automatically change the style of the model response depending on who is asking the question. The education provider will give the model the age range of the user who has asked the question.
Which solution meets these requirements with the LEAST implementation effort?
A. Fine-tune the model by using additional training data that is representative of the various age ranges that the application will support.
B. Summarize the response text depending on the age of the user so that younger users receive shorter responses.
C. Add a role description to the prompt context that instructs the model of the age range that the response should target.
D. Use chain-of-thought reasoning to deduce the correct style and complexity for a response suitable for that user.
Answer: C (LEAVE A REPLY)

NEW QUESTION: 67
A media company wants to analyze viewer behavior and demographics to recommend personalized content.
The company wants to deploy a customized ML model in its production environment. The company also wants to observe if the model quality drifts over time.
Which AWS service or feature meets these requirements?
A. Amazon Rekognition
B. Amazon SageMaker Clarify
C. Amazon Comprehend
D. Amazon SageMaker Model Monitor
Answer: D (LEAVE A REPLY)

The requirement is to deploy a customized machine learning (ML) model and monitor its quality for potential drift over time in a production environment. Let's evaluate each option:
* A. Amazon Rekognition: This service is designed for image and video analysis, such as object detection, facial recognition, and text extraction. It is not suited for deploying custom ML models or monitoring model quality drift.
* B. Amazon SageMaker Clarify: This feature helps detect bias in ML models and explains model predictions. While it addresses fairness and interpretability, it does not specifically focus on monitoring model quality drift over time in production.
* C. Amazon Comprehend: This is a natural language processing (NLP) service for extracting insights from text, such as sentiment analysis or entity recognition. It does not support deploying custom ML models or monitoring model performance drift.
* D. Amazon SageMaker Model Monitor: This feature is part of Amazon SageMaker and is specifically designed to monitor ML models in production. It tracks metrics such as data drift, model drift, and performance degradation over time, alerting users when issues are detected.
Exact Extract Reference: According to the AWS documentation on Amazon SageMaker, "Amazon SageMaker Model Monitor allows you to detect and remediate data and model quality issues in production. It continuously monitors the performance of deployed models, capturing data and model predictions to detect deviations from expected behavior, such as data drift or model performance degradation." (Source: AWS SageMaker Documentation - Model Monitoring, https://docs.aws.amazon.com/sagemaker/latest/dg/model-monitor.html).
This directly aligns with the requirement to observe model quality drift, making Amazon SageMaker Model Monitor the correct choice.
:
AWS SageMaker Documentation: Model Monitoring
(https://docs.aws.amazon.com/sagemaker/latest/dg
/model-monitor.html)
AWS AI Practitioner Study Guide (conceptual alignment with monitoring deployed ML models)

NEW QUESTION: 68
A company needs to choose a model from Amazon Bedrock to use internally. The company must identify a model that generates responses in a style that the company's employees prefer.
What should the company do to meet these requirements?
A. Use public model leaderboards to identify the model.
B. Use the model InvocationLatency runtime metrics in Amazon CloudWatch when trying models.
C. Evaluate the models by using built-in prompt datasets.
D. Evaluate the models by using a human workforce and custom prompt datasets.

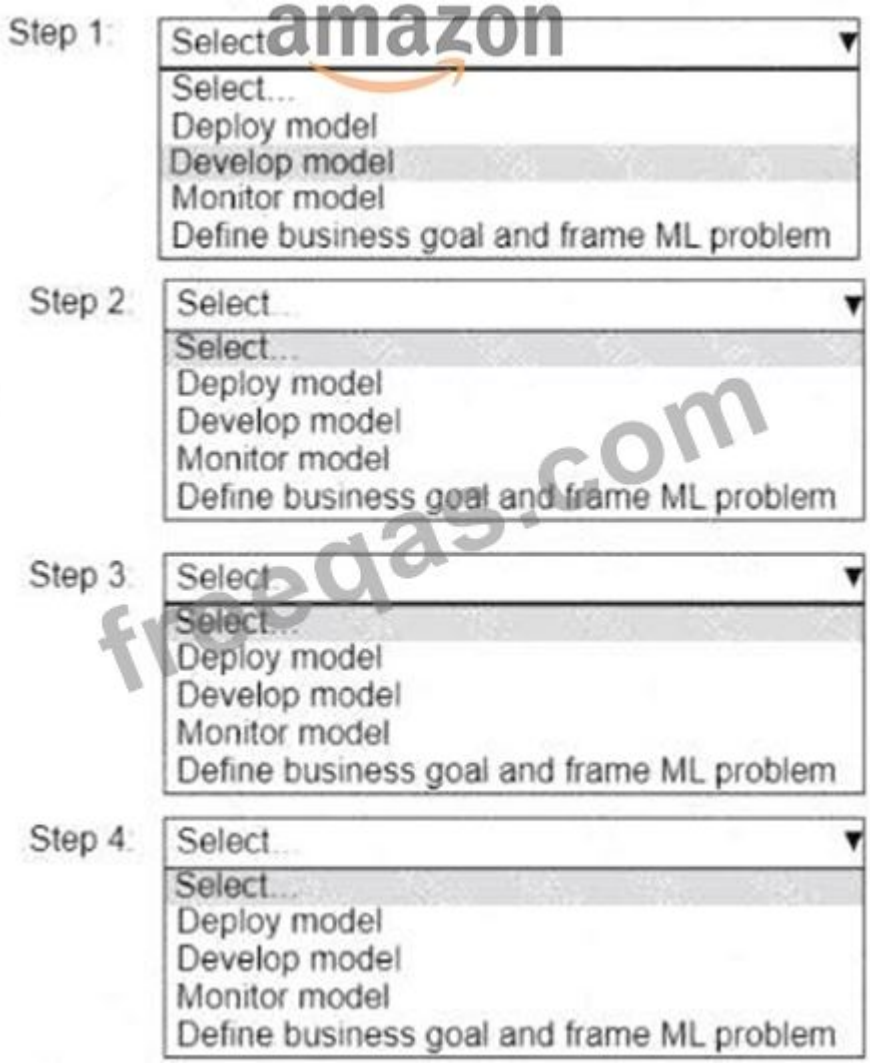
Answer: D (LEAVE A REPLY)

NEW QUESTION: 69

A company wants to build an ML application.

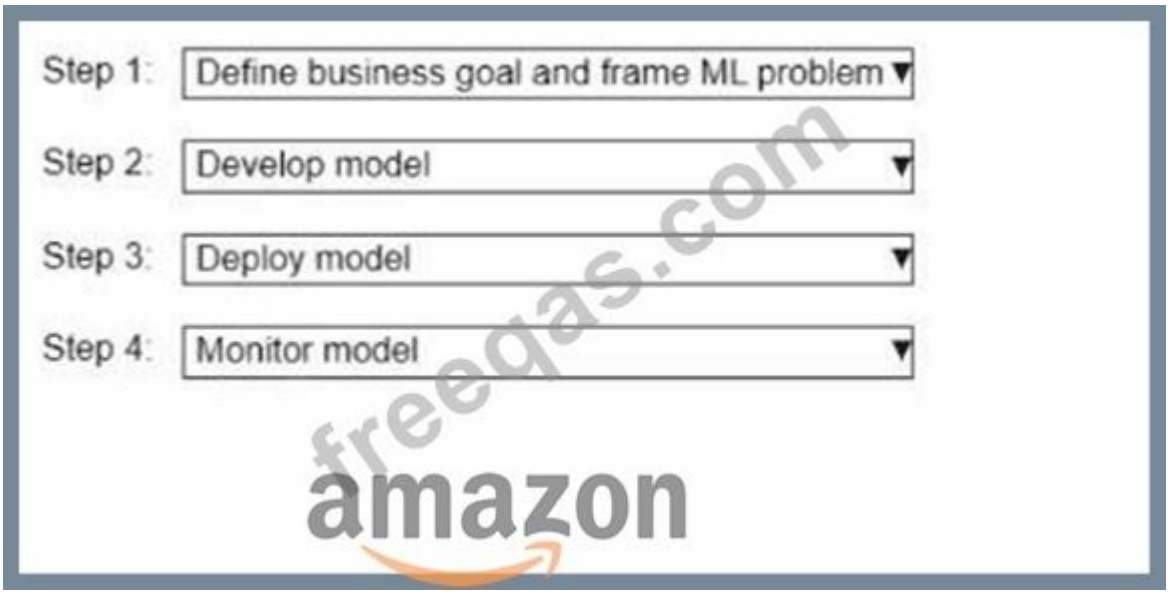
Select and order the correct steps from the following list to develop a well-architected ML workload. Each step should be selected one time. (Select and order FOUR.)

* Deploy model
* Develop model
* Monitor model
* Define business goal and frame ML problem



Answer:

Explanation:



Building a well-architected ML workload follows a structured lifecycle as outlined in AWS best practices.

The process begins with defining the business goal and framing the ML problem to ensure the project aligns with organizational objectives. Next, the model is developed, which includes data preparation, training, and evaluation. Once the model is ready, it is deployed tomake predictions in a production environment. Finally, the model is monitored to ensure it performs as expected and to address any issues like drift or degradation over time. This order ensures a systematic approach to ML development.

Exact Extract from AWS AI Documents:

From the AWS AI Practitioner Learning Path:

"The machine learning lifecycle typically follows these stages: 1) Define the business goal and frame the ML problem, 2) Develop the model (including data preparation, training, and evaluation), 3) Deploy the model to production, and 4) Monitor the model for performance and drift to ensure it continues to meet business needs." (Source: AWS AI Practitioner Learning Path, Module on Machine Learning Lifecycle) Detailed Explanation:

* Step 1: Define business goal and frame ML problemThis is the first step in any ML project. It involves understanding the business objective (e.g., reducing churn) and framing the ML problem (e.g., classification or regression). Without this step, the project lacks direction. The hotspot lists this option as "Define business goal and frame ML problem," which matches this stage.
* Step 2: Develop modelAfter defining the problem, the next step is to develop the model. This includes collecting and preparing data, selecting an algorithm, training the model, and evaluating its performance. The hotspot lists "Develop model" as an option, aligning with this stage.
* Step 3: Deploy modelOnce the model is developed and meets performance requirements, it is deployed to a production environment to make predictions or automate decisions. The hotspot includes "Deploy model" as an option, which fits this stage.
* Step 4: Monitor modelAfter deployment, the model must be monitored to ensure it performs well over time, addressing issues like data drift or performance degradation. The hotspot lists "Monitor model" as an option, completing the lifecycle.
Hotspot Selection Analysis:
The hotspot provides four steps, each with the same dropdown options: "Select...," "Deploy model," "Develop model," "Monitor model," and "Define business goal and frame ML problem." The correct selections are:
* Step 1: Define business goal and frame ML problem
* Step 2: Develop model
* Step 3: Deploy model
* Step 4: Monitor model
Each option is used exactly once, as required, and follows the logical order of the ML lifecycle.
References:
AWS AI Practitioner Learning Path: Module on Machine Learning Lifecycle Amazon SageMaker Developer Guide: Machine Learning Workflow (https://docs.aws.amazon.com/sagemaker/latest/dg/how-it-works-mlconcepts.html)
AWS Well-Architected Framework: Machine Learning Lens (https://docs.aws.amazon.com/wellarchitected/latest/machine-learning-lens/)

NEW QUESTION: 70
Which scenario represents a practical use case for generative AI?
A. Using an ML model to forecast product demand
B. Employing a chatbot to provide human-like responses to customer queries in real time
C. Using an analytics dashboard to track website traffic and user behavior
D. Implementing a rule-based recommendation engine to suggest products to customers
Answer: B (LEAVE A REPLY)

Generative AI is a type of AI that creates new content, such as text, images, or audio, often mimicking human- like outputs. A practical use case for generative AI is employing a chatbot to provide human-like responses to customer queries in real time, as it leverages the ability of large language models (LLMs) to generate natural language responses dynamically.
Exact Extract from AWS AI Documents:
From the AWS Bedrock User Guide:
"Generative AI enables applications like chatbots to produce human-like text responses in real time, enhancing customer support by providing natural and contextually relevant answers to user queries." (Source: AWS Bedrock User Guide, Introduction to Generative AI) Detailed Explanation:
* Option A: Using an ML model to forecast product demandForecasting product demand typically involves predictive analytics using supervised learning (e.g., regression models), not generative AI, which focuses on creating new content.
* Option B: Employing a chatbot to provide human-like responses to customer queries in real time This is the correct answer. Generative AI, particularly LLMs, is commonly used to power chatbots that generate human-like responses, making this a practical use case.
* Option C: Using an analytics dashboard to track website traffic and user behaviorAn analytics dashboard involves data visualization and analysis, not generative AI, which is about creating new content.
* Option D: Implementing a rule-based recommendation engine to suggest products to customersA rule-based recommendation engine relies on predefined rules, not generative AI. Generative AI could be used for more dynamic recommendations, but this scenario does not describe such a case.
References:
AWS Bedrock User Guide: Introduction to Generative AI
(https://docs.aws.amazon.com/bedrock/latest
/userguide/what-is-bedrock.html)
AWS AI Practitioner Learning Path: Module on Generative AI Applications AWS Documentation: Generative AI Use Cases (https://aws.amazon.com/generative-ai/)

NEW QUESTION: 71
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).

Answer: A (LEAVE A REPLY)

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 References:

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.

NEW QUESTION: 72
A company is training a foundation model (FM). The company wants to increase the accuracy of the model up to a specific acceptance level.
Which solution will meet these requirements?
A. Increase the epochs.
B. Decrease the epochs.
C. Increase the temperature parameter.
D. Decrease the batch size.
Answer: A (LEAVE A REPLY)

NEW QUESTION: 73
Which component of Amazon Bedrock Studio can help secure the content that AI systems generate?
A. Access controls
B. Function calling
C. Guardrails
D. Knowledge bases
Answer: (SHOW ANSWER)
Amazon Bedrock Studio provides tools to build and manage generative AI applications, and the company needs a component to secure the content generated by AI systems. Guardrails in Amazon Bedrock are designed to ensure safe and responsible AI outputs by filtering harmful or inappropriate content, making them the key component for securing generated content.
Exact Extract from AWS AI Documents:
From the AWS Bedrock User Guide:
"Guardrails in Amazon Bedrock provide mechanisms to secure the content generated by AI systems by filtering out harmful or inappropriate outputs, such as hate speech, violence, or misinformation, ensuring responsible AI usage." (Source: AWS Bedrock User Guide, Guardrails for Responsible AI) Detailed Explanation:
* Option A: Access controlsAccess controls manage who can use or interact with the AI system but do not directly secure the content generated by the system.
* Option B: Function callingFunction calling enables AI models to interact with external tools or APIs, but it is not related to securing generated content.
* Option C: GuardrailsThis is the correct answer. Guardrails in Amazon Bedrock secure generated content by filtering out harmful or inappropriate material, ensuring safe outputs.
* Option D: Knowledge basesKnowledge bases provide data for AI models to generate responses but do not inherently secure the content that is generated.
References:

AWS Bedrock User Guide: Guardrails for Responsible AI (https://docs.aws.amazon.com/bedrock/latest /userguide/guardrails.html)
AWS AI Practitioner Learning Path: Module on Responsible AI and Model Safety Amazon Bedrock Developer Guide: Securing AI Outputs (https://aws.amazon.com/bedrock/)

NEW QUESTION: 74
A company deployed an AI/ML solution to help customer service agents respond to frequently asked questions. The questions can change over time. The company wants to give customer service agents the ability to ask questions and receive automatically generated answers to common customer questions.
Which strategy will meet these requirements MOST cost-effectively?
A. Fine-tune the model regularly.
B. Train the model by using context data.
C. Pre-train and benchmark the model by using context data.
D. Use Retrieval Augmented Generation (RAG) with prompt engineering techniques.
Answer: D (LEAVE A REPLY)
RAG combines large pre-trained models with retrieval mechanisms to fetch relevant context from a knowledge base. This approach is cost-effective as it eliminates the need for frequent model retraining while ensuring responses are contextually accurate and up to date. References: AWS RAG Techniques.

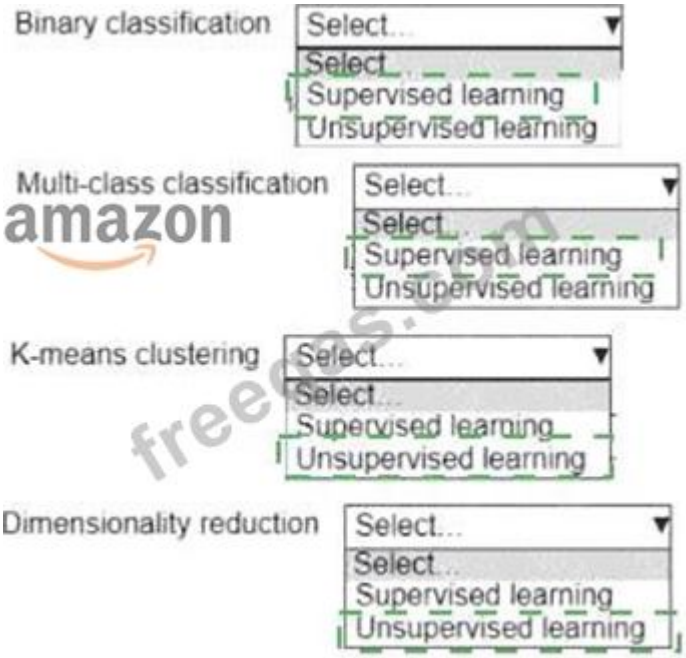
NEW QUESTION: 75
A company wants to make a chatbot to help customers. The chatbot will help solve technical problems without human intervention. The company chose a foundation model (FM) for the chatbot. The chatbot needs to produce responses that adhere to company tone.
Which solution meets these requirements?
A. Set a low limit on the number of tokens the FM can produce.
B. Use batch inferencing to process detailed responses.
C. Experiment and refine the prompt until the FM produces the desired responses.
D. Define a higher number for the temperature parameter.
Answer: C (LEAVE A REPLY)
Experimenting and refining the prompt is the best approach to ensure that the chatbot using a foundation model (FM) produces responses that adhere to the company's tone.
* Prompt Engineering:
* Adjusting and refining the prompt allows for better control over the FM's outputs, ensuring they align with the desired tone and style.
* This iterative process involves testing different prompts and modifying them based on the model's responses to achieve the desired outcome.
* Why Option C is Correct:

* Directly Influences Output Quality: Allows for fine-tuning of the model's responses to match the company's tone.
* Cost-Effective: Does not require modifying the model itself, only the inputs to it.
* Why Other Options are Incorrect:
* A. Low limit on tokens: Limits response length but not the adherence to company tone.
* B. Batch inferencing: Relates to processing multiple inputs, not controlling response tone.
* D. Higher temperature: Increases randomness in responses, which could deviate from the desired tone.

NEW QUESTION: 76
Which option is a use case for generative AI models?
A. Improving network security by using intrusion detection systems
B. Enhancing database performance by using optimized indexing
C. Creating photorealistic images from text descriptions for digital marketing
D. Analyzing financial data to forecast stock market trends
Answer: C (LEAVE A REPLY)

NEW QUESTION: 77
A company wants to develop ML applications to improve business operations and efficiency.
Select the correct ML paradigm from the following list for each use case. Each ML paradigm should be selected one or more times. (Select FOUR.)
* Supervised learning
* Unsupervised learning

Answer:



Explanation:

The company is developing ML applications for various use cases, and the task is to select the correct ML paradigm (supervised or unsupervised learning) for each. Supervised learning involves training a model on labeled data to make predictions, while unsupervised

learning identifies patterns or structures in unlabeled data. Each use case aligns with one of these paradigms based on its requirements.
Exact Extract from AWS AI Documents:
From the AWS AI Practitioner Learning Path:
"Supervised learning uses labeled data to train models for tasks like classification (e.g., binary or multi-class classification), where the model predicts a category. Unsupervised learning works with unlabeled data for tasks like clustering (e.g., K-means clustering) or dimensionality reduction, identifying patternsor reducing data complexity without predefined labels." (Source: AWS AI Practitioner Learning Path, Module on Machine Learning Strategies) Detailed Explanation:
Binary classification: Supervised learningBinary classification involves predicting one of two classes (e.g., yes
/no, spam/not spam) using labeled data, making it a supervised learning task. The model learns from examples where the correct class is provided.
Multi-class classification: Supervised learningMulti-class classification extends binary classification to predict one of multiple classes (e.g., categorizing items into several groups). Like binary classification, it requires labeled data, so it falls under supervised learning.
K-means clustering: Unsupervised learningK-means clustering groups data into clusters based on similarity, without requiring labeled data. This is a classic unsupervised learning task, as the algorithm identifies patterns in the data on its own.
Dimensionality reduction: Unsupervised learningDimensionality reduction (e.g., using techniques like PCA) reduces the number of features in a dataset while preserving important information. It does not require labeled data, making it an unsupervised learning task.
Hotspot Selection Analysis:
The hotspot lists four use cases, each with a dropdown containing "Select...," "Supervised learning," and
"Unsupervised learning." The correct selections are:
Binary classification: Supervised learning
Multi-class classification: Supervised learning
K-means clustering: Unsupervised learning
Dimensionality reduction: Unsupervised learning
Each paradigm (supervised and unsupervised learning) is used twice, as the question allows for paradigms to be selected one or more times.
References:
AWS AI Practitioner Learning Path: Module on Machine Learning Strategies Amazon SageMaker Developer Guide: Supervised and Unsupervised Learning (https://docs.aws.amazon.com
/sagemaker/latest/dg/algos.html)

AWS Documentation: Introduction to Machine Learning Paradigms (https://aws.amazon.com/machine- learning/)

NEW QUESTION: 78
A company wants to use a large language model (LLM) to develop a conversational agent. The company needs to prevent the LLM from being manipulated with common prompt engineering techniques to perform undesirable actions or expose sensitive information.
Which action will reduce these risks?
A. Create a prompt template that teaches the LLM to detect attack patterns.
B. Increase the temperature parameter on invocation requests to the LLM.
C. Avoid using LLMs that are not listed in Amazon SageMaker.
D. Decrease the number of input tokens on invocations of the LLM.
Answer: (SHOW ANSWER)
Creating a prompt template that teaches the LLM to detect attack patterns is the most effective way to reduce the risk of the model being manipulated through prompt engineering.
* Prompt Templates for Security:
* A well-designed prompt template can guide the LLM to recognize and respond appropriately to potential manipulation attempts.
* This strategy helps prevent the model from performing undesirable actions or exposing sensitive information by embedding security awareness directly into the prompts.
* Why Option A is Correct:
* Teaches Model Security Awareness: Equips the LLM to handle potentially harmful inputs by recognizing suspicious patterns.
* Reduces Manipulation Risk: Helps mitigate risks associated with prompt engineering attacks by proactively preparing the LLM.
* Why Other Options are Incorrect:
* B. Increase the temperature parameter: This increases randomness in responses, potentially making the LLM more unpredictable and less secure.
* C. Avoid LLMs not listed in SageMaker: Does not directly address the risk of prompt manipulation.
* D. Decrease the number of input tokens: Does not mitigate risks related to prompt manipulation.

NEW QUESTION: 79
A company wants to use a large language model (LLM) on Amazon Bedrock for sentiment analysis. The company needs the LLM to produce more consistent responses to the same input prompt.
Which adjustment to an inference parameter should the company make to meet these requirements?
A. Decrease the temperature value

B. Increase the temperature value
C. Decrease the length of output tokens
D. Increase the maximum generation length
Answer: (SHOW ANSWER)
The temperature parameter in a large language model (LLM) controls the randomness of the model's output.
A lower temperature value makes the output more deterministic and consistent, meaning that the model is less likely to produce different results for the same input prompt.
* Option A (Correct): "Decrease the temperature value": This is the correct answer because lowering the temperature reduces the randomness of the responses, leading to more consistent outputs for the same input.
* Option B: "Increase the temperature value" is incorrect because it would make the output more random and less consistent.
* Option C: "Decrease the length of output tokens" is incorrect as it does not directly affect the consistency of the responses.
* Option D: "Increase the maximum generation length" is incorrect because this adjustment affects the output length, not the consistency of the model's responses.
AWS AI Practitioner References:
* Understanding Temperature in Generative AI Models: AWS documentation explains that adjusting the temperature parameter affects the model's output randomness, with lower values providing more consistent outputs.

NEW QUESTION: 80
A financial institution is building an AI solution to make loan approval decisions by using a foundation model (FM). For security and audit purposes, the company needs the AI solution's decisions to be explainable.
Which factor relates to the explainability of the AI solution's decisions?
A. Model complexity
B. Training time
C. Number of hyperparameters
D. Deployment time
Answer: A (LEAVE A REPLY)
The financial institution needs an AI solution for loan approval decisions to be explainable for security and audit purposes. Explainability refers to the ability to understand and interpret how a model makes decisions.
Model complexity directly impacts explainability: simpler models (e.g., logistic regression) are more interpretable, while complex models (e.g., deep neural networks) are harder to explain, often behaving like
"black boxes."
Exact Extract from AWS AI Documents:
From the Amazon SageMaker Developer Guide:

"Model complexity affects the explainability of AI solutions. Simpler models, such as linear regression, are inherently more interpretable, while complex models, such as deep neural networks, may require additional tools like SageMaker Clarify to provide insights into their decision-making processes." (Source: Amazon SageMaker Developer Guide, Explainability with SageMaker Clarify) Detailed Explanation:
* Option A: Model complexityThis is the correct answer. The complexity of the model directly influences how easily its decisions can be explained, a critical factor for audit and security purposes in loan approvals.
* Option B: Training timeTraining time refers to how long it takes to train the model, which does not directly impact the explainability of its decisions.
* Option C: Number of hyperparametersWhile hyperparameters affect model performance, they do not directly relate to explainability. A model with many hyperparameters might still be explainable if it is a simple model.
* Option D: Deployment timeDeployment time refers to the time taken to deploy the model to production, which is unrelated to the explainability of its decisions.
References:
Amazon SageMaker Developer Guide: Explainability with SageMaker Clarify (https://docs.aws.amazon.com /sagemaker/latest/dg/clarify-explainability.html)
AWS AI Practitioner Learning Path: Module on Responsible AI and Explainability AWS Documentation: Explainable AI (https://aws.amazon.com/machine-learning/responsible-ai/)

NEW QUESTION: 81
An AI practitioner is using an Amazon Bedrock base model to summarize session chats from the customer service department. The AI practitioner wants to store invocation logs to monitor model input and output data.
Which strategy should the AI practitioner use?
A. Enable invocation logging in Amazon Bedrock.
B. Configure AWS Audit Manager as the logs destination for the model.
C. Configure model invocation logging in Amazon EventBridge.
D. Configure AWS CloudTrail as the logs destination for the model.
Answer: A (LEAVE A REPLY)

NEW QUESTION: 82
A company wants to deploy a conversational chatbot to answer customer questions. The chatbot is based on a fine-tuned Amazon SageMaker JumpStart model. The application must comply with multiple regulatory frameworks.
Which capabilities can the company show compliance for? (Select TWO.)
A. Auto scaling inference endpoints
B. Threat detection
C. Data protection

D. Cost optimization
E. Loosely coupled microservices
Answer: (SHOW ANSWER)
Let me know if you'd like to continue with any more questions or if you need further assistance!

NEW QUESTION: 83
A company deployed a model to production. After 4 months, the model inference quality degraded. The company wants to receive a notification if the model inference quality degrades. The company also wants to ensure that the problem does not happen again. Which solution will meet these requirements?
A. Retrain the model. Monitor model drift by using Amazon SageMaker Clarify.
B. Retrain the model. Monitor model drift by using Amazon SageMaker Model Monitor.
C. Build a new model. Monitor model drift by using Amazon SageMaker Feature Store.
D. Build a new model. Monitor model drift by using Amazon SageMaker JumpStart.
Answer: B (LEAVE A REPLY)
The company needs to address the degradation in model inference quality after 4 months in production and prevent future occurrences by receiving notifications. Retraining the model can address the current degradation, likely caused by data drift (changes in the data distribution over time). Amazon SageMaker Model Monitor is designed to detect and monitor model drift, alerting the company when inference quality degrades, thus meeting both requirements.
Exact Extract from AWS AI Documents:
From the Amazon SageMaker Developer Guide:
"Amazon SageMaker Model Monitor enables you to monitor machine learning models in production for data drift, model performance degradation, and other quality issues. It can detect drift in feature distributions and inference quality, sending notifications when deviations are detected, allowing you to take corrective actions such as retraining the model." (Source: Amazon SageMaker Developer Guide, Monitoring Models with SageMaker Model Monitor) Detailed Explanation:
* Option A: Retrain the model. Monitor model drift by using Amazon SageMaker Clarify.
SageMaker Clarify is used for bias detection and explainability, not for monitoring model drift or inference quality in production. This option does not fully meet the requirements.
* Option B: Retrain the model. Monitor model drift by using Amazon SageMaker Model Monitor.
This is the correct answer. Retraining addresses the current degradation, and SageMaker Model Monitor can detect future drift in inference quality, sending notifications to prevent recurrence, as required.
* Option C: Build a new model. Monitor model drift by using Amazon SageMaker Feature Store.

SageMaker Feature Store is for managing and sharing features, not for monitoring model drift or inference quality. Building a new model may not be necessary if retraining can address the issue.
* Option D: Build a new model. Monitor model drift by using Amazon SageMaker JumpStart.
SageMaker JumpStart provides pre-trained models and solutions for quick deployment, but it does not offer specific tools for monitoring model drift or inference quality in production.
References:
Amazon SageMaker Developer Guide: Monitoring Models with SageMaker Model Monitor (https://docs.aws.
amazon.com/sagemaker/latest/dg/model-monitor.html)
AWS AI Practitioner Learning Path: Module on Model Monitoring and Maintenance AWS Documentation: Addressing Model Drift in Production
(https://aws.amazon.com/sagemaker/)

NEW QUESTION: 84
Which option describes embeddings in the context of AI?
A. A method for compressing large datasets
B. An encryption method for securing sensitive data
C. A method for visualizing high-dimensional data
D. A numerical method for data representation in a reduced dimensionality space
Answer: D (LEAVE A REPLY)
Embeddings in AI refer to numerical representations of data (e.g., text, images) in a lower-dimensional space, capturing semantic or contextual relationships. They are widely used in NLP and other AI tasks to represent complex data in a format that models can process efficiently.
Exact Extract from AWS AI Documents:
From the AWS AI Practitioner Learning Path:
"Embeddings are numerical representations of data in a reduced dimensionality space. In natural language processing, for example, word or sentence embeddings capture semantic relationships, enabling models to process text efficiently for tasks like classification or similarity search." (Source: AWS AI Practitioner Learning Path, Module on AI Concepts)
Detailed Explanation:
* Option A: A method for compressing large datasetsWhile embeddings reduce dimensionality, their primary purpose is not data compression but rather to represent data in a way that preserves meaningful relationships. This option is incorrect.
* Option B: An encryption method for securing sensitive dataEmbeddings are not related to encryption or data security. They are used for data representation, making this option incorrect.

* Option C: A method for visualizing high-dimensional dataWhile embeddings can sometimes be used in visualization (e.g., t-SNE), their primary role is data representation for model processing, not visualization. This option is misleading.
* Option D: A numerical method for data representation in a reduced dimensionality spaceThis is the correct answer. Embeddings transform complex data into lower-dimensional numerical vectors, preserving semantic or contextual information for use in AI models.
References:
AWS AI Practitioner Learning Path: Module on AI Concepts
Amazon Comprehend Developer Guide: Embeddings for Text Analysis (https://docs.aws.amazon.com /comprehend/latest/dg/embeddings.html)
AWS Documentation: What are Embeddings? (https://aws.amazon.com/what-is/embeddings/)

NEW QUESTION: 85
An AI practitioner has a database of animal photos. The AI practitioner wants to automatically identify and categorize the animals in the photos without manual human effort.
Which strategy meets these requirements?
A. Anomaly detection
B. Object detection
C. Inpainting
D. Named entity recognition
Answer: B (LEAVE A REPLY)

NEW QUESTION: 86
A company wants to create an application by using Amazon Bedrock. The company has a limited budget and prefers flexibility without long-term commitment.
Which Amazon Bedrock pricing model meets these requirements?
A. Spot Instance
B. Provisioned Throughput
C. On-Demand
D. Model customization
Answer: C (LEAVE A REPLY)

NEW QUESTION: 87
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)
Answer: (SHOW ANSWER)
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.