

# Amazon-Web-Services

## Exam Questions AIF-C01

AWS Certified AI Practitioner



#### NEW QUESTION 1

A company wants to use a large language model (LLM) on Amazon Bedrock for sentiment analysis. The company wants to know how much information can fit into one prompt.

Which consideration will inform the company's decision?

- A. Temperature
- B. Context window
- C. Batch size
- D. Model size

**Answer: B**

#### Explanation:

The context window determines how much information can fit into a single prompt when using a large language model (LLM) like those on Amazon Bedrock.

? Context Window:

? Why Option B is Correct:

? Why Other Options are Incorrect:

#### NEW QUESTION 2

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. Configure AWS CloudTrail as the logs destination for the model.
- B. Enable invocation logging in Amazon Bedrock.
- C. Configure AWS Audit Manager as the logs destination for the model.
- D. Configure model invocation logging in Amazon EventBridge.

**Answer: B**

#### Explanation:

Amazon Bedrock provides an option to enable invocation logging to capture and store the input and output data of the models used. This is essential for monitoring and auditing purposes, particularly when handling customer data.

? Option B (Correct): "Enable invocation logging in Amazon Bedrock": This is the correct answer as it directly enables the logging of all model invocations, ensuring transparency and traceability.

? Option A: "Configure AWS CloudTrail" is incorrect because CloudTrail logs API calls but does not provide specific logging for model inputs and outputs.

? Option C: "Configure AWS Audit Manager" is incorrect as Audit Manager is used for compliance reporting, not specific invocation logging for AI models.

? Option D: "Configure model invocation logging in Amazon EventBridge" is incorrect as EventBridge is for event-driven architectures, not specifically designed for logging AI model inputs and outputs.

AWS AI Practitioner References:

? Amazon Bedrock Logging Capabilities: AWS emphasizes using built-in logging features in Bedrock to maintain data integrity and transparency in model operations.

#### NEW QUESTION 3

A company is building a large language model (LLM) question answering chatbot. The company wants to decrease the number of actions call center employees need to take to respond to customer questions.

Which business objective should the company use to evaluate the effect of the LLM chatbot?

- A. Website engagement rate
- B. Average call duration
- C. Corporate social responsibility
- D. Regulatory compliance

**Answer: B**

#### Explanation:

The business objective to evaluate the effect of an LLM chatbot aimed at reducing the actions required by call center employees should be average call duration.

? Average Call Duration:

? Why Option B is Correct:

? Why Other Options are Incorrect:

#### NEW QUESTION 4

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 EC2
- B. Amazon Q Developer
- C. Amazon Q in Amazon QuickSight
- D. Amazon Q in AWS Chatbot

**Answer: C**

#### Explanation:

Amazon QuickSight is a fully managed business intelligence (BI) service that allows users to create and publish interactive dashboards that include visualizations like graphs, charts, and tables. "Amazon Q" is the natural language query feature within Amazon QuickSight. It enables users to ask questions about their data in natural language and receive visual responses such as graphs.

? Option C (Correct): "Amazon Q in Amazon QuickSight": This is the correct answer

because Amazon QuickSight Q is specifically designed to allow users to explore their data through natural language queries, and it can automatically generate graphs to display sales data and other metrics. This makes it an ideal choice for the company to automate the generation of graphs showing total sales for its top-selling products across various retail locations.

? Option A, B, and D: These options are incorrect:

AWS AI Practitioner References:

? Amazon QuickSight Q is designed to provide insights from data by using natural language queries, making it a powerful tool for generating automated graphs and visualizations directly from queried data.

? Business Intelligence (BI) on AWS: AWS services such as Amazon QuickSight provide business intelligence capabilities, including automated reporting and visualization features, which are ideal for companies seeking to visualize data like sales trends over time.

#### NEW QUESTION 5

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

#### Explanation:

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:

? Why Option C is Correct:

? Why Other Options are Incorrect:

#### NEW QUESTION 6

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

#### Explanation:

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:

? Why Option A is Correct:

? Why Other Options are Incorrect:

#### NEW QUESTION 7

A company has a foundation model (FM) that was customized by using Amazon Bedrock to answer customer queries about products. The company wants to validate the model's responses to new types of queries. The company needs to upload a new dataset that Amazon Bedrock can use for validation.

Which AWS service meets these requirements?

- A. Amazon S3
- B. Amazon Elastic Block Store (Amazon EBS)
- C. Amazon Elastic File System (Amazon EFS)
- D. AWS Snowcone

**Answer:** A

#### Explanation:

Amazon S3 is the optimal choice for storing and uploading datasets used for machine learning model validation and training. It offers scalable, durable, and secure storage, making it ideal for holding datasets required by Amazon Bedrock for validation purposes.

? Option A (Correct): "Amazon S3": This is the correct answer because Amazon S3

is widely used for storing large datasets that are accessed by machine learning models, including those in Amazon Bedrock.

? Option B: "Amazon Elastic Block Store (Amazon EBS)" is incorrect because EBS

is a block storage service for use with Amazon EC2, not for directly storing datasets for Amazon Bedrock.

? Option C: "Amazon Elastic File System (Amazon EFS)" is incorrect as it is

primarily used for file storage with shared access by multiple instances.

? Option D: "AWS Snowcone" is incorrect because it is a physical device for offline data transfer, not suitable for directly providing data to Amazon Bedrock.

AWS AI Practitioner References:

? Storing and Managing Datasets on AWS for Machine Learning: AWS recommends using S3 for storing and managing datasets required for ML model training and validation.

#### NEW QUESTION 8

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**

**Explanation:**

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:

? Why Option B is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 9**

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**

**Explanation:**

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 10**

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**

**Explanation:**

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 10**

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**

**Explanation:**

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 12**

A company built a deep learning model for object detection and deployed the model to production.

Which AI process occurs when the model analyzes a new image to identify objects?

- A. Training
- B. Inference
- C. Model deployment
- D. Bias correction

**Answer: B**

**Explanation:**

Inference is the correct answer because it is the AI process that occurs when a deployed model analyzes new data (such as an image) to make predictions or identify objects.

? Inference:

? Why Option B is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 13**

A company wants to develop a large language model (LLM) application by using Amazon Bedrock and customer data that is uploaded to Amazon S3. The company's security policy states that each team can access data for only the team's own customers.

Which solution will meet these requirements?

- A. Create an Amazon Bedrock custom service role for each team that has access to only the team's customer data.
- B. Create a custom service role that has Amazon S3 access
- C. Ask teams to specify the customer name on each Amazon Bedrock request.
- D. Redact personal data in Amazon S3. Update the S3 bucket policy to allow team access to customer data.
- E. Create one Amazon Bedrock role that has full Amazon S3 access
- F. Create IAM roles for each team that have access to only each team's customer folders.

**Answer: A**

**Explanation:**

To comply with the company's security policy, which restricts each team to access data for only their own customers, creating an Amazon Bedrock custom service role for each team is the correct solution.

? Custom Service Role Per Team:

? Why Option A is Correct:

? Why Other Options are Incorrect:

Thus, A is the correct answer to meet the company's security requirements.

**NEW QUESTION 14**

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**

**Explanation:**

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 16**

A company wants to use a pre-trained generative AI model to generate content for its marketing campaigns. The company needs to ensure that the generated content aligns with the company's brand voice and messaging requirements.

Which solution meets these requirements?

- A. Optimize the model's architecture and hyperparameters to improve the model's overall performance.

- B. Increase the model's complexity by adding more layers to the model's architecture.
- C. Create effective prompts that provide clear instructions and context to guide the model's generation.
- D. Select a large, diverse dataset to pre-train a new generative model.

**Answer: C**

**Explanation:**

Creating effective prompts is the best solution to ensure that the content generated by a pre-trained generative AI model aligns with the company's brand voice and messaging requirements.

? Effective Prompt Engineering:

? Why Option C is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 17**

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. Use Amazon Macie to monitor SageMaker Studio.
- C. Configure SageMaker to use a VPC with an S3 endpoint.
- D. Configure SageMaker to use S3 Glacier Deep Archive.

**Answer: C**

**Explanation:**

To manage the flow of data from Amazon S3 to SageMaker Studio notebooks securely, using a VPC with an S3 endpoint is the best solution.

? Amazon SageMaker and S3 Integration:

? Why Option C is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 19**

Which metric measures the runtime efficiency of operating AI models?

- A. Customer satisfaction score (CSAT)
- B. Training time for each epoch
- C. Average response time
- D. Number of training instances

**Answer: C**

**Explanation:**

The average response time is the correct metric for measuring the runtime efficiency of operating AI models.

? Average Response Time:

? Why Option C is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 24**

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. Reinforcement learning with rewards for positive customer feedback
- C. Unsupervised learning to find clusters of similar customer inquiries
- D. Supervised learning with a continuously updated FAQ database

**Answer: B**

**Explanation:**

Reinforcement learning allows a model to learn and improve over time based on feedback from its environment. In this case, the chatbot can improve its responses by being rewarded for positive customer feedback, which aligns well with the goal of self-improvement based on past interactions and new information.

? Option B (Correct): "Reinforcement learning with rewards for positive customer feedback": This is the correct answer as reinforcement learning enables the chatbot to learn from feedback and adapt its behavior accordingly, providing self-improvement capabilities.

? Option A: "Supervised learning with a manually curated dataset" is incorrect because it does not support continuous learning from new interactions.

? Option C: "Unsupervised learning to find clusters of similar customer inquiries" is incorrect because unsupervised learning does not provide a mechanism for improving responses based on feedback.

? Option D: "Supervised learning with a continuously updated FAQ database" is incorrect because it still relies on manually curated data rather than self-improvement from feedback.

AWS AI Practitioner References:

? Reinforcement Learning on AWS: AWS provides reinforcement learning frameworks that can be used to train models to improve their performance based on feedback.

**NEW QUESTION 29**

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. On-Demand
- B. Model customization
- C. Provisioned Throughput
- D. Spot Instance

**Answer:** A

**Explanation:**

Amazon Bedrock offers an on-demand pricing model that provides flexibility without long-term commitments. This model allows companies to pay only for the resources they use, which is ideal for a limited budget and offers flexibility.

? Option A (Correct): "On-Demand": This is the correct answer because on-demand pricing allows the company to use Amazon Bedrock without any long-term commitments and to manage costs according to their budget.

? Option B: "Model customization" is a feature, not a pricing model.

? Option C: "Provisioned Throughput" involves reserving capacity ahead of time, which might not offer the desired flexibility and could lead to higher costs if the capacity is not fully used.

? Option D: "Spot Instance" is a pricing model for EC2 instances and does not apply to Amazon Bedrock.

AWS AI Practitioner References:

? AWS Pricing Models for Flexibility: On-demand pricing is a key AWS model for services that require flexibility and no long-term commitment, ensuring cost-effectiveness for projects with variable usage patterns.

**NEW QUESTION 30**

What are tokens in the context of generative AI models?

- A. Tokens are the basic units of input and output that a generative AI model operates on, representing words, subwords, or other linguistic units.
- B. Tokens are the mathematical representations of words or concepts used in generative AI models.
- C. Tokens are the pre-trained weights of a generative AI model that are fine-tuned for specific tasks.
- D. Tokens are the specific prompts or instructions given to a generative AI model to generate output.

**Answer:** A

**Explanation:**

Tokens in generative AI models are the smallest units that the model processes, typically representing words, subwords, or characters. They are essential for the model to understand and generate language, breaking down text into manageable parts for processing.

? Option A (Correct): "Tokens are the basic units of input and output that a

generative AI model operates on, representing words, subwords, or other linguistic units": This is the correct definition of tokens in the context of generative AI models.

? Option B: "Mathematical representations of words" describes embeddings, not tokens.

? Option C: "Pre-trained weights of a model" refers to the parameters of a model, not tokens.

? Option D: "Prompts or instructions given to a model" refers to the queries or commands provided to a model, not tokens.

AWS AI Practitioner References:

? Understanding Tokens in NLP: AWS provides detailed explanations of how tokens are used in natural language processing tasks by AI models, such as in Amazon Comprehend and other AWS AI services.

**NEW QUESTION 32**

A social media company wants to use a large language model (LLM) for content moderation. The company wants to evaluate the LLM outputs for bias and potential discrimination against specific groups or individuals.

Which data source should the company use to evaluate the LLM outputs with the LEAST administrative effort?

- A. User-generated content
- B. Moderation logs
- C. Content moderation guidelines
- D. Benchmark datasets

**Answer:** D

**Explanation:**

Benchmark datasets are pre-validated datasets specifically designed to evaluate machine learning models for bias, fairness, and potential discrimination. These datasets are the most efficient tool for assessing an LLM's performance against known standards with minimal administrative effort.

? Option D (Correct): "Benchmark datasets": This is the correct answer because using standardized benchmark datasets allows the company to evaluate model outputs for bias with minimal administrative overhead.

? Option A: "User-generated content" is incorrect because it is unstructured and would require significant effort to analyze for bias.

? Option B: "Moderation logs" is incorrect because they represent historical data and do not provide a standardized basis for evaluating bias.

? Option C: "Content moderation guidelines" is incorrect because they provide qualitative criteria rather than a quantitative basis for evaluation.

AWS AI Practitioner References:

? Evaluating AI Models for Bias on AWS: AWS supports using benchmark datasets to assess model fairness and detect potential bias efficiently.

**NEW QUESTION 34**

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

**Answer:** A

**Explanation:**

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

Explanation of Each Option:

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

? Option B: "Mask the confidential data in the inference responses by using dynamic data masking." This option is incorrect because dynamic data masking is typically used to mask or obfuscate sensitive data in a database. It does not address the core problem of the model being trained 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 38**

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:** A

**Explanation:**

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 42**

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
- E. Use Amazon Rekognition to add custom labels to the data.

**Answer:** B

**Explanation:**

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:

? Why Option B is Correct:

? Why Other Options are Incorrect:

Thus, B is the correct answer for meeting transparency and explainability requirements for the foundation model

**NEW QUESTION 43**

Which option is a use case for generative AI models?

- A. Improving network security by using intrusion detection systems
- B. Creating photorealistic images from text descriptions for digital marketing

- C. Enhancing database performance by using optimized indexing
- D. Analyzing financial data to forecast stock market trends

**Answer:** B

**Explanation:**

Generative AI models are used to create new content based on existing data. One common use case is generating photorealistic images from text descriptions, which is particularly useful in digital marketing, where visual content is key to engaging potential customers.

? Option B (Correct): "Creating photorealistic images from text descriptions for digital marketing": This is the correct answer because generative AI models, like those offered by Amazon Bedrock, can create images based on text descriptions, making them highly valuable for generating marketing materials.

? Option A: "Improving network security by using intrusion detection systems" is incorrect because this is a use case for traditional machine learning models, not generative AI.

? Option C: "Enhancing database performance by using optimized indexing" is incorrect as it is unrelated to generative AI.

? Option D: "Analyzing financial data to forecast stock market trends" is incorrect because it typically involves predictive modeling rather than generative AI.

AWS AI Practitioner References:

? Use Cases for Generative AI Models on AWS: AWS highlights the use of generative AI for creative content generation, including image creation, text generation, and more, which is suited for digital marketing applications.

**NEW QUESTION 48**

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:** A

**Explanation:**

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:

? Why Option A is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 52**

A company manually reviews all submitted resumes in PDF format. As the company grows, the company expects the volume of resumes to exceed the company's review capacity. The company needs an automated system to convert the PDF resumes into plain text format for additional processing.

Which AWS service meets this requirement?

- A. Amazon Textract
- B. Amazon Personalize
- C. Amazon Lex
- D. Amazon Transcribe

**Answer:** A

**Explanation:**

Amazon Textract is a service that automatically extracts text and data from scanned documents, including PDFs. It is the best choice for converting resumes from PDF format to plain text for further processing.

? Amazon Textract:

? Why Option A is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 54**

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

**Explanation:**

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 56

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:** D

#### Explanation:

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 59

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. AWS PrivateLink
- B. Amazon Macie
- C. Amazon CloudFront
- D. Internet gateway

**Answer:** A

#### Explanation:

AWS PrivateLink enables private connectivity between VPCs and AWS services without exposing traffic to the public internet. This feature is critical for meeting regulatory compliance standards that require isolation from public internet traffic.

? Option A (Correct): "AWS PrivateLink": This is the correct answer because it allows secure access to Amazon Bedrock and other AWS services from a VPC without internet access, ensuring compliance with regulatory standards.

? Option B: "Amazon Macie" is incorrect because it is a security service for data classification and protection, not for managing private network traffic.

? Option C: "Amazon CloudFront" is incorrect because it is a content delivery network service and does not provide private network connectivity.

? Option D: "Internet gateway" is incorrect as it enables internet access, which violates the VPC's no-internet-traffic policy.

AWS AI Practitioner References:

? AWS PrivateLink Documentation: AWS highlights PrivateLink as a solution for connecting VPCs to AWS services privately, which is essential for organizations with strict regulatory requirements.

#### NEW QUESTION 61

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. Generative pre-trained transformers (GPT)
- B. Residual neural network
- C. Support vector machine
- D. WaveNet

**Answer:** A

#### Explanation:

Generative Pre-trained Transformers (GPT) are suitable for building an AI-based application that can generate SQL queries from natural language input provided by employees.

? GPT for Natural Language Processing:

? Why Option A is Correct:

? Why Other Options are Incorrect:

#### NEW QUESTION 63

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:** BD

**Explanation:**

To determine customer sentiments based on written customer reviews, the company can use Amazon Comprehend and Amazon Bedrock.

? Amazon Comprehend:

? Amazon Bedrock:

? Why Other Options are Incorrect:

**NEW QUESTION 67**

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. Build a conversational chatbot by using Amazon Lex.
- B. Transcribe call recordings by using Amazon Transcribe.
- C. Extract information from call recordings by using Amazon SageMaker Model Monitor.
- D. Create classification labels by using Amazon Comprehend.

**Answer:** B

**Explanation:**

Amazon Transcribe is the correct solution for converting audio from customer calls into text, allowing the company to analyze and extract key information from the conversations.

? Amazon Transcribe:

? Why Option B is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 70**

A company is building a chatbot to improve user experience. The company is using a large language model (LLM) from Amazon Bedrock for intent detection. The company wants to use few-shot learning to improve intent detection accuracy.

Which additional data does the company need to meet these requirements?

- A. Pairs of chatbot responses and correct user intents
- B. Pairs of user messages and correct chatbot responses
- C. Pairs of user messages and correct user intents
- D. Pairs of user intents and correct chatbot responses

**Answer:** C

**Explanation:**

Few-shot learning involves providing a model with a few examples (shots) to learn from. For improving intent detection accuracy in a chatbot using a large language model (LLM), the data should consist of pairs of user messages and their corresponding correct intents.

? Few-shot Learning for Intent Detection:

? Why Option C is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 72**

A company wants to develop an educational game where users answer questions such as the following: "A jar contains six red, four green, and three yellow marbles. What is the probability of choosing a green marble from the jar?"

Which solution meets these requirements with the LEAST operational overhead?

- A. Use supervised learning to create a regression model that will predict probability.
- B. Use reinforcement learning to train a model to return the probability.
- C. Use code that will calculate probability by using simple rules and computations.
- D. Use unsupervised learning to create a model that will estimate probability density.

**Answer:** C

**Explanation:**

The problem involves a simple probability calculation that can be handled efficiently by straightforward mathematical rules and computations. Using machine learning techniques would introduce unnecessary complexity and operational overhead.

? Option C (Correct): "Use code that will calculate probability by using simple rules and computations": This is the correct answer because it directly solves the problem with minimal overhead, using basic probability rules.

? Option A: "Use supervised learning to create a regression model" is incorrect as it overcomplicates the solution for a simple probability problem.

? Option B: "Use reinforcement learning to train a model" is incorrect because reinforcement learning is not needed for a simple probability calculation.

? Option D: "Use unsupervised learning to create a model" is incorrect as unsupervised learning is not applicable to this task.

AWS AI Practitioner References:

? Choosing the Right Solution for AI Tasks: AWS recommends using the simplest and most efficient approach to solve a given problem, avoiding unnecessary machine learning techniques for straightforward tasks.

**NEW QUESTION 76**

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. Toxicity
- B. Hallucinations
- C. Plagiarism
- D. Privacy

**Answer:** C

**Explanation:**

The scenario where a student copies content from generative AI to write essays represents the challenge of plagiarism in responsible AI use.

? Plagiarism:

? Why Option C is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 77**

A company has built a solution by using generative AI. The solution uses large language models (LLMs) to translate training manuals from English into other languages. The company wants to evaluate the accuracy of the solution by examining the text generated for the manuals.

Which model evaluation strategy meets these requirements?

- A. Bilingual Evaluation Understudy (BLEU)
- B. Root mean squared error (RMSE)
- C. Recall-Oriented Understudy for Gisting Evaluation (ROUGE)
- D. F1 score

**Answer:** A

**Explanation:**

BLEU (Bilingual Evaluation Understudy) is a metric used to evaluate the accuracy of machine-generated translations by comparing them against reference translations. It is commonly used for translation tasks to measure how close the generated output is to professional human translations.

? Option A (Correct): "Bilingual Evaluation Understudy (BLEU)": This is the correct answer because BLEU is specifically designed to evaluate the quality of translations, making it suitable for the company's use case.

? Option B: "Root mean squared error (RMSE)" is incorrect because RMSE is used for regression tasks to measure prediction errors, not translation quality.

? Option C: "Recall-Oriented Understudy for Gisting Evaluation (ROUGE)" is incorrect as it is used to evaluate text summarization, not translation.

? Option D: "F1 score" is incorrect because it is typically used for classification tasks, not for evaluating translation accuracy.

AWS AI Practitioner References:

? Model Evaluation Metrics on AWS: AWS supports various metrics like BLEU for specific use cases, such as evaluating machine translation models.

**NEW QUESTION 81**

A pharmaceutical company wants to analyze user reviews of new medications and provide a concise overview for each medication. Which solution meets these requirements?

- A. Create a time-series forecasting model to analyze the medication reviews by using Amazon Personalize.
- B. Create medication review summaries by using Amazon Bedrock large language models (LLMs).
- C. Create a classification model that categorizes medications into different groups by using Amazon SageMaker.
- D. Create medication review summaries by using Amazon Rekognition.

**Answer:** B

**Explanation:**

Amazon Bedrock provides large language models (LLMs) that are optimized for natural language understanding and text summarization tasks, making it the best choice for creating concise summaries of user reviews. Time-series forecasting, classification, and image analysis (Rekognition) are not suitable for summarizing textual data. References: AWS Bedrock Documentation.

**NEW QUESTION 85**

A company is building a solution to generate images for protective eyewear. The solution must have high accuracy and must minimize the risk of incorrect annotations.

Which solution will meet these requirements?

- A. Human-in-the-loop validation by using Amazon SageMaker Ground Truth Plus
- B. Data augmentation by using an Amazon Bedrock knowledge base
- C. Image recognition by using Amazon Rekognition
- D. Data summarization by using Amazon QuickSight

**Answer:** A

**Explanation:**

Amazon SageMaker Ground Truth Plus is a managed data labeling service that includes human-in-the-loop (HITL) validation. This solution ensures high accuracy by involving human reviewers to validate the annotations and reduce the risk of incorrect annotations.

? Amazon SageMaker Ground Truth Plus:

? Why Option A is Correct:

? Why Other Options are Incorrect:

Thus, A is the correct answer for generating high-accuracy images with minimized annotation risks.

**NEW QUESTION 86**

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:** AC

**Explanation:**

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

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

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

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

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

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

AWS AI Practitioner References:

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

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

**NEW QUESTION 89**

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:** C

**Explanation:**

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 90**

A company is building an application that needs to generate synthetic data that is based on existing data. Which type of model can the company use to meet this requirement?

- A. Generative adversarial network (GAN)
- B. XGBoost
- C. Residual neural network
- D. WaveNet

**Answer:** A

**Explanation:**

Generative adversarial networks (GANs) are a type of deep learning model used for generating synthetic data based on existing datasets. GANs consist of two neural networks (a generator and a discriminator) that work together to create realistic data.

? Option A (Correct): "Generative adversarial network (GAN)": This is the correct answer because GANs are specifically designed for generating synthetic data that closely resembles the real data they are trained on.

? Option B: "XGBoost" is a gradient boosting algorithm for classification and regression tasks, not for generating synthetic data.

? Option C: "Residual neural network" is primarily used for improving the performance of deep networks, not for generating synthetic data.

? Option D: "WaveNet" is a model architecture designed for generating raw audio waveforms, not synthetic data in general.

AWS AI Practitioner References:

? GANs on AWS for Synthetic Data Generation: AWS supports the use of GANs for creating synthetic datasets, which can be crucial for applications like training machine learning models in environments where real data is scarce or sensitive.

**NEW QUESTION 95**

A company wants to create a chatbot by using a foundation model (FM) on Amazon Bedrock. The FM needs to access encrypted data that is stored in an Amazon S3 bucket. The data is encrypted with Amazon S3 managed keys (SSE-S3). The FM encounters a failure when attempting to access the S3 bucket data. Which solution will meet these requirements?

- A. Ensure that the role that Amazon Bedrock assumes has permission to decrypt data with the correct encryption key.
- B. Set the access permissions for the S3 buckets to allow public access to enable access over the internet.
- C. Use prompt engineering techniques to tell the model to look for information in Amazon S3.
- D. Ensure that the S3 data does not contain sensitive information.

**Answer:** A

**Explanation:**

Amazon Bedrock needs the appropriate IAM role with permission to access and decrypt data stored in Amazon S3. If the data is encrypted with Amazon S3 managed keys (SSE- S3), the role that Amazon Bedrock assumes must have the required permissions to access and decrypt the encrypted data.

? Option A (Correct): "Ensure that the role that Amazon Bedrock assumes has permission to decrypt data with the correct encryption key": This is the correct solution as it ensures that the AI model can access the encrypted data securely without changing the encryption settings or compromising data security.

? Option B: "Set the access permissions for the S3 buckets to allow public access" is incorrect because it violates security best practices by exposing sensitive data to the public.

? Option C: "Use prompt engineering techniques to tell the model to look for information in Amazon S3" is incorrect as it does not address the encryption and permission issue.

? Option D: "Ensure that the S3 data does not contain sensitive information" is incorrect because it does not solve the access problem related to encryption.

AWS AI Practitioner References:

? Managing Access to Encrypted Data in AWS: AWS recommends using proper IAM roles and policies to control access to encrypted data stored in S3.

**NEW QUESTION 99**

.....

## **Thank You for Trying Our Product**

### **We offer two products:**

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

### **AIF-C01 Practice Exam Features:**

- \* AIF-C01 Questions and Answers Updated Frequently
- \* AIF-C01 Practice Questions Verified by Expert Senior Certified Staff
- \* AIF-C01 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- \* AIF-C01 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

**100% Actual & Verified — Instant Download, Please Click**  
**[Order The AIF-C01 Practice Test Here](#)**