

# Amazon-Web-Services

## Exam Questions AIF-C01

AWS Certified AI Practitioner



### NEW QUESTION 1

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 2

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 3

A law firm wants to build an AI application by using large language models (LLMs). The application will read legal documents and extract key points from the documents. Which solution meets these requirements?

- A. Build an automatic named entity recognition system.
- B. Create a recommendation engine.
- C. Develop a summarization chatbot.
- D. Develop a multi-language translation system.

**Answer:** C

#### Explanation:

A summarization chatbot is ideal for extracting key points from legal documents. Large language models (LLMs) can be used to summarize complex texts, such as legal documents, making them more accessible and understandable.

? Option C (Correct): "Develop a summarization chatbot": This is the correct answer

because a summarization chatbot uses LLMs to condense and extract key information from text, which is precisely the requirement for reading and summarizing legal documents.

? Option A: "Build an automatic named entity recognition system" is incorrect

because it focuses on identifying specific entities, not summarizing documents.

? Option B: "Create a recommendation engine" is incorrect as it is used to suggest products or content, not summarize text.

? Option D: "Develop a multi-language translation system" is incorrect because translation is unrelated to summarizing text.

AWS AI Practitioner References:

? Using LLMs for Text Summarization on AWS: AWS supports developing summarization tools using its AI services, including Amazon Bedrock.

### NEW QUESTION 4

An AI practitioner has built a deep learning model to classify the types of materials in images. The AI practitioner now wants to measure the model performance. Which metric will help the AI practitioner evaluate the performance of the model?

- A. Confusion matrix
- B. Correlation matrix
- C. R2 score
- D. Mean squared error (MSE)

**Answer:** A

**Explanation:**

A confusion matrix is the correct metric for evaluating the performance of a classification model, such as the deep learning model built to classify types of materials in images.

? Confusion Matrix:

? Why Option A is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 5**

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. Topic modeling
- B. Clustering models
- C. Prescriptive ML models
- D. BERT-based models

**Answer:** D

**Explanation:**

BERT-based models (Bidirectional Encoder Representations from Transformers) are suitable for tasks that involve understanding the context of words in a sentence and suggesting missing words. These models use bidirectional training, which considers the context from both directions (left and right of the missing word) to predict the appropriate word to fill in the gaps.

? BERT-based Models:

? Why Option D is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 6**

A company uses a foundation model (FM) from Amazon Bedrock for an AI search tool. The company wants to fine-tune the model to be more accurate by using the company's data.

Which strategy will successfully fine-tune the model?

- A. Provide labeled data with the prompt field and the completion field.
- B. Prepare the training dataset by creating a .txt file that contains multiple lines in .csv format.
- C. Purchase Provisioned Throughput for Amazon Bedrock.
- D. Train the model on journals and textbooks.

**Answer:** A

**Explanation:**

Providing labeled data with both a prompt field and a completion field is the correct strategy for fine-tuning a foundation model (FM) on Amazon Bedrock.

? Fine-Tuning Strategy:

? 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 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

**Explanation:**

Amazon SageMaker JumpStart is the correct service for quickly deploying and consuming a foundation model (FM) within a team's VPC.

? Amazon SageMaker JumpStart:

? Why Option B is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 9**

A company built an AI-powered resume screening system. The company used a large dataset to train the model. The dataset contained resumes that were not representative of all demographics. Which core dimension of responsible AI does this scenario present?

- A. Fairness.
- B. Explainability.
- C. Privacy and security.
- D. Transparency.

**Answer:** A

**Explanation:**

Fairness refers to the absence of bias in AI models. Using non-representative datasets leads to biased predictions, affecting specific demographics unfairly. Explainability, privacy, and transparency are important but not directly related to this scenario. References: AWS Responsible AI Framework.

**NEW QUESTION 10**

Which term describes the numerical representations of real-world objects and concepts that AI and natural language processing (NLP) models use to improve understanding of textual information?

- A. Embeddings
- B. Tokens
- C. Models
- D. Binaries

**Answer:** A

**Explanation:**

Embeddings are numerical representations of objects (such as words, sentences, or documents) that capture the objects' semantic meanings in a form that AI and NLP models can easily understand. These representations help models improve their understanding of textual information by representing concepts in a continuous vector space.

? Option A (Correct): "Embeddings": This is the correct term, as embeddings provide

a way for models to learn relationships between different objects in their input space, improving their understanding and processing capabilities.

? Option B: "Tokens" are pieces of text used in processing, but they do not capture semantic meanings like embeddings do.

? Option C: "Models" are the algorithms that use embeddings and other inputs, not the representations themselves.

? Option D: "Binaries" refer to data represented in binary form, which is unrelated to the concept of embeddings.

AWS AI Practitioner References:

? Understanding Embeddings in AI and NLP: AWS provides resources and tools, like Amazon SageMaker, that utilize embeddings to represent data in formats suitable for machine learning models.

**NEW QUESTION 10**

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

**Explanation:**

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 11**

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 12**

A company wants to use large language models (LLMs) with Amazon Bedrock to develop a chat interface for the company's product manuals. The manuals are stored as PDF files.

Which solution meets these requirements MOST cost-effectively?

- A. Use prompt engineering to add one PDF file as context to the user prompt when the prompt is submitted to Amazon Bedrock.
- B. Use prompt engineering to add all the PDF files as context to the user prompt when the prompt is submitted to Amazon Bedrock.
- C. Use all the PDF documents to fine-tune a model with Amazon Bedrock.
- D. Use the fine-tuned model to process user prompts.
- E. Upload PDF documents to an Amazon Bedrock knowledge base.
- F. Use the knowledge base to provide context when users submit prompts to Amazon Bedrock.

**Answer: A**

**Explanation:**

Using Amazon Bedrock with large language models (LLMs) allows for efficient utilization of AI to answer queries based on context provided in product manuals. To achieve this cost-effectively, the company should avoid unnecessary use of resources.

? Option A (Correct): "Use prompt engineering to add one PDF file as context to the user prompt when the prompt is submitted to Amazon Bedrock": This is the most cost-effective solution. By using prompt engineering, only the relevant content from one PDF file is added as context to each query. This approach minimizes the amount of data processed, which helps in reducing costs associated with LLMs' computational requirements.

? Option B: "Use prompt engineering to add all the PDF files as context to the user prompt when the prompt is submitted to Amazon Bedrock" is incorrect. Including all PDF files would increase costs significantly due to the large context size processed by the model.

? Option C: "Use all the PDF documents to fine-tune a model with Amazon Bedrock" is incorrect. Fine-tuning a model is more expensive than using prompt engineering, especially if done for multiple documents.

? Option D: "Upload PDF documents to an Amazon Bedrock knowledge base" is incorrect because Amazon Bedrock does not have a built-in knowledge base feature for directly managing and querying PDF documents.

AWS AI Practitioner References:

? Prompt Engineering for Cost-Effective AI: AWS emphasizes the importance of using prompt engineering to minimize costs when interacting with LLMs. By carefully selecting relevant context, users can reduce the amount of data processed and save on expenses.

**NEW QUESTION 16**

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: A**

**Explanation:**

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.

Reference: Amazon SageMaker Model Cards

**NEW QUESTION 18**

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 23

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 24

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 27

An AI company periodically evaluates its systems and processes with the help of independent software vendors (ISVs). The company needs to receive email message notifications when an ISV's compliance reports become available. Which AWS service can the company use to meet this requirement?

- A. AWS Audit Manager
- B. AWS Artifact
- C. AWS Trusted Advisor
- D. AWS Data Exchange

**Answer:** D

#### Explanation:

AWS Data Exchange is a service that allows companies to securely exchange data with third parties, such as independent software vendors (ISVs). AWS Data Exchange can be configured to provide notifications, including email notifications, when new datasets or compliance reports become available.

? Option D (Correct): "AWS Data Exchange": This is the correct answer because it enables the company to receive notifications, including email messages, when ISVs' compliance reports are available.

? Option A: "AWS Audit Manager" is incorrect because it focuses on assessing an organization's own compliance, not receiving third-party compliance reports.

? Option B: "AWS Artifact" is incorrect as it provides access to AWS's compliance reports, not ISVs'.

? Option C: "AWS Trusted Advisor" is incorrect as it offers optimization and best practices guidance, not compliance report notifications.

AWS AI Practitioner References:

? AWS Data Exchange Documentation: AWS explains how Data Exchange allows organizations to subscribe to third-party data and receive notifications when updates are available.

### NEW QUESTION 31

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:** A

**Explanation:**

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:

? Why Option A is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 32**

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:** AC

**Explanation:**

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:

? C. Modify the Training Data to Mitigate Bias:

? Why Other Options are Incorrect:

**NEW QUESTION 37**

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 39**

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 42**

A company has installed a security camera. The company uses an ML model to evaluate the security camera footage for potential thefts. The company has discovered that the model disproportionately flags people who are members of a specific ethnic group.

Which type of bias is affecting the model output?

- A. Measurement bias
- B. Sampling bias
- C. Observer bias
- D. Confirmation bias

**Answer:** B

**Explanation:**

Sampling bias is the correct type of bias affecting the model output when it disproportionately flags people from a specific ethnic group.

? Sampling Bias:

? Why Option B is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 45**

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 46**

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

**Explanation:**

Object detection is the correct strategy for automatically identifying and categorizing animals in photos.

? Object Detection:

? Why Option A is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 47**

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 49**

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 Wrangle
- C. Create ML models and demand forecast predictions by using SageMaker built-in algorithms.
- D. Import the data into Amazon SageMaker Data Wrangle
- E. Build ML models and demand forecast predictions by using an Amazon Personalize Trending-Now recipe.
- F. Import the data into Amazon SageMaker Canvas
- G. Build ML models and demand forecast predictions by selecting the values in the data from SageMaker Canvas.

**Answer:** D

**Explanation:**

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 53**

A company has a database of petabytes of unstructured data from internal sources. The company wants to transform this data into a structured format so that its data scientists can perform machine learning (ML) tasks.

Which service will meet these requirements?

- A. Amazon Lex
- B. Amazon Rekognition
- C. Amazon Kinesis Data Streams
- D. AWS Glue

**Answer:** D

**Explanation:**

AWS Glue is the correct service for transforming petabytes of unstructured data into a structured format suitable for machine learning tasks.

? AWS Glue:

? Why Option D is Correct:

? Why Other Options are Incorrect:

**NEW QUESTION 56**

Which feature of Amazon OpenSearch Service gives companies the ability to build vector database applications?

- A. Integration with Amazon S3 for object storage
- B. Support for geospatial indexing and queries
- C. Scalable index management and nearest neighbor search capability
- D. Ability to perform real-time analysis on streaming data

**Answer:** C

**Explanation:**

Amazon OpenSearch Service (formerly Amazon Elasticsearch Service) has introduced capabilities to support vector search, which allows companies to build vector database applications. This is particularly useful in machine learning, where vector representations (embeddings) of data are often used to capture semantic meaning.

Scalable index management and nearest neighbor search capability are the core features enabling vector database functionalities in OpenSearch. The service allows users to index high-dimensional vectors and perform efficient nearest neighbor searches, which are crucial for tasks such as recommendation systems, anomaly detection, and semantic search.

Here is why option C is the correct Answer:

? Scalable Index Management: OpenSearch Service supports scalable indexing of vector data. This means you can index a large volume of high-dimensional vectors

and manage these indexes in a cost-effective and performance-optimized way. The service leverages underlying AWS infrastructure to ensure that indexing scales seamlessly with data size.

? Nearest Neighbor Search Capability: OpenSearch Service's nearest neighbor search capability allows for fast and efficient searches over vector data. This is essential for applications like product recommendation engines, where the system needs to quickly find the most similar items based on a user's query or behavior.

? AWS AI Practitioner References:

The other options do not directly relate to building vector database applications:

? A. Integration with Amazon S3 for object storage is about storing data objects, not vector-based searching or indexing.

? B. Support for geospatial indexing and queries is related to location-based data, not vectors used in machine learning.

? D. Ability to perform real-time analysis on streaming data relates to analyzing incoming data streams, which is different from the vector search capabilities.

**NEW QUESTION 60**

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. Data pre-processing
- B. Feature engineering
- C. Exploratory data analysis
- D. Hyperparameter tuning

**Answer:** C

**Explanation:**

Exploratory data analysis (EDA) involves understanding the data by visualizing it, calculating statistics, and creating correlation matrices. This stage helps identify patterns, relationships, and anomalies in the data, which can guide further steps in the ML pipeline.

? Option C (Correct): "Exploratory data analysis": This is the correct answer as the tasks described (correlation matrix, calculating statistics, visualizing data) are all part of the EDA process.

? Option A: "Data pre-processing" is incorrect because it involves cleaning and transforming data, not initial analysis.

? Option B: "Feature engineering" is incorrect because it involves creating new features from raw data, not analyzing the data's existing structure.

? Option D: "Hyperparameter tuning" is incorrect because it refers to optimizing model parameters, not analyzing the data.

AWS AI Practitioner References:

? Stages of the Machine Learning Pipeline: AWS outlines EDA as the initial phase of understanding and exploring data before moving to more specific preprocessing, feature engineering, and model training stages.

**NEW QUESTION 61**

A company wants to build an interactive application for children that generates new stories based on classic stories. The company wants to use Amazon Bedrock and needs to ensure that the results and topics are appropriate for children.

Which AWS service or feature will meet these requirements?

- A. Amazon Rekognition
- B. Amazon Bedrock playgrounds
- C. Guardrails for Amazon Bedrock
- D. Agents for Amazon Bedrock

**Answer:** C

**Explanation:**

Amazon Bedrock is a service that provides foundational models for building generative AI applications. When creating an application for children, it is crucial to ensure that the generated content is appropriate for the target audience. "Guardrails" in Amazon Bedrock provide mechanisms to control the outputs and topics of generated content to align with desired safety standards and appropriateness levels.

? Option C (Correct): "Guardrails for Amazon Bedrock": This is the correct answer because guardrails are specifically designed to help users enforce content moderation, filtering, and safety checks on the outputs generated by models in Amazon Bedrock. For a children's application, guardrails ensure that all content generated is suitable and appropriate for the intended audience.

? Option A: "Amazon Rekognition" is incorrect. Amazon Rekognition is an image and video analysis service that can detect inappropriate content in images or videos, but it does not handle text or story generation.

? Option B: "Amazon Bedrock playgrounds" is incorrect because playgrounds are environments for experimenting and testing model outputs, but they do not inherently provide safeguards to ensure content appropriateness for specific audiences, such as children.

? Option D: "Agents for Amazon Bedrock" is incorrect. Agents in Amazon Bedrock facilitate building AI applications with more interactive capabilities, but they do not provide specific guardrails for ensuring content appropriateness for children.

AWS AI Practitioner References:

? Guardrails in Amazon Bedrock: Designed to help implement controls that ensure generated content is safe and suitable for specific use cases or audiences, such as children, by moderating and filtering inappropriate or undesired content.

? Building Safe AI Applications: AWS provides guidance on implementing ethical AI practices, including using guardrails to protect against generating inappropriate or biased content.

**NEW QUESTION 66**

A company uses Amazon SageMaker for its ML pipeline in a production environment. The company has large input data sizes up to 1 GB and processing times up to 1 hour. The company needs near real-time latency.

Which SageMaker inference option meets these requirements?

- A. Real-time inference
- B. Serverless inference
- C. Asynchronous inference
- D. Batch transform

**Answer:** A

**Explanation:**

Real-time inference is designed to provide immediate, low-latency predictions, which is necessary when the company requires near real-time latency for its ML models. This option is optimal when there is a need for fast responses, even with large input data sizes and substantial processing times.

? Option A (Correct): "Real-time inference": This is the correct answer because it supports low-latency requirements, which are essential for real-time applications where quick response times are needed.

? Option B: "Serverless inference" is incorrect because it is more suited for intermittent, small-scale inference workloads, not for continuous, large-scale, low-latency needs.

? Option C: "Asynchronous inference" is incorrect because it is used for workloads that do not require immediate responses.

? Option D: "Batch transform" is incorrect as it is intended for offline, large-batch processing where immediate response is not necessary.

AWS AI Practitioner References:

? Amazon SageMaker Inference Options: AWS documentation describes real-time inference as the best solution for applications that require immediate prediction results with low latency.

**NEW QUESTION 67**

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 70**

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 71**

A company is building an ML model to analyze archived data. The company must perform inference on large datasets that are multiple GBs in size. The company does not need to access the model predictions immediately.

Which Amazon SageMaker inference option will meet these requirements?

- A. Batch transform
- B. Real-time inference
- C. Serverless inference
- D. Asynchronous inference

**Answer:** A

**Explanation:**

Batch transform in Amazon SageMaker is designed for offline processing of large datasets. It is ideal for scenarios where immediate predictions are not required, and the inference can be done on large datasets that are multiple gigabytes in size. This method processes data in batches, making it suitable for analyzing archived data without the need for real-time access to predictions.

? Option A (Correct): "Batch transform": This is the correct answer because batch transform is optimized for handling large datasets and is suitable when immediate access to predictions is not required.

? Option B: "Real-time inference" is incorrect because it is used for low-latency, real-time prediction needs, which is not required in this case.

? Option C: "Serverless inference" is incorrect because it is designed for small-scale, intermittent inference requests, not for large batch processing.

? Option D: "Asynchronous inference" is incorrect because it is used when immediate predictions are required, but with high throughput, whereas batch transform is more suitable for very large datasets.

AWS AI Practitioner References:

? Batch Transform on AWS SageMaker: AWS recommends using batch transform for large datasets when real-time processing is not needed, ensuring cost-effectiveness and scalability.

**NEW QUESTION 76**

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

**Explanation:**

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:

? Why Option C is Correct:

? Why Other Options are Incorrect:

Thus, C is the correct answer for detecting suspicious IP addresses.

#### NEW QUESTION 78

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 80

A company is using domain-specific models. The company wants to avoid creating new models from the beginning. The company instead wants to adapt pre-trained models to create models for new, related tasks.

Which ML strategy meets these requirements?

- A. Increase the number of epochs.
- B. Use transfer learning.
- C. Decrease the number of epochs.
- D. Use unsupervised learning.

**Answer: B**

#### Explanation:

Transfer learning is the correct strategy for adapting pre-trained models for new, related tasks without creating models from scratch.

? Transfer Learning:

? Why Option B is Correct:

? Why Other Options are Incorrect:

#### NEW QUESTION 82

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**

#### Explanation:

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 83

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 86**

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 89**

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 91**

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 95**

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 96

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 100

Which option is a benefit of using Amazon SageMaker Model Cards to document AI models?

- A. Providing a visually appealing summary of a model's capabilities.
- B. Standardizing information about a model's purpose, performance, and limitations.
- C. Reducing the overall computational requirements of a model.
- D. Physically storing models for archival purposes.

**Answer: B**

#### Explanation:

Amazon SageMaker Model Cards provide a standardized way to document important details about an AI model, such as its purpose, performance, intended usage, and known limitations. This enables transparency and compliance while fostering better communication between stakeholders. It does not store models physically or optimize computational requirements. References: AWS SageMaker Model Cards Documentation.

#### NEW QUESTION 101

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. Decrease the batch size.
- B. Increase the epochs.
- C. Decrease the epochs.
- D. Increase the temperature parameter.

**Answer: B**

#### Explanation:

Increasing the number of epochs during model training allows the model to learn from the data over more iterations, potentially improving its accuracy up to a certain point. This is a common practice when attempting to reach a specific level of accuracy.

? Option B (Correct): "Increase the epochs": This is the correct answer because increasing epochs allows the model to learn more from the data, which can lead to higher accuracy.

? Option A: "Decrease the batch size" is incorrect as it mainly affects training speed and may lead to overfitting but does not directly relate to achieving a specific accuracy level.

? Option C: "Decrease the epochs" is incorrect as it would reduce the training time, possibly preventing the model from reaching the desired accuracy.

? Option D: "Increase the temperature parameter" is incorrect because temperature affects the randomness of predictions, not model accuracy.

AWS AI Practitioner References:

? Model Training Best Practices on AWS: AWS suggests adjusting training parameters, like the number of epochs, to improve model performance.

**NEW QUESTION 104**

.....

## **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](#)**