



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 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 3

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 4

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 5

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 6

A company needs to choose a model from Amazon Bedrock to use internally. The company must identify a model that generates responses in a style that the company's employees prefer.

What should the company do to meet these requirements?

- A. Evaluate the models by using built-in prompt datasets.
- B. Evaluate the models by using a human workforce and custom prompt datasets.
- C. Use public model leaderboards to identify the model.
- D. Use the model InvocationLatency runtime metrics in Amazon CloudWatch when trying models.

Answer: B

Explanation:

To determine which model generates responses in a style that the company's employees prefer, the best approach is to use a human workforce to evaluate the models with custom prompt datasets. This method allows for subjective evaluation based on the specific stylistic preferences of the company's employees, which cannot be effectively assessed through automated methods or pre-built datasets.

? Option B (Correct): "Evaluate the models by using a human workforce and custom

prompt datasets": This is the correct answer as it directly involves human judgment to evaluate the style and quality of the responses, aligning with employee preferences.

? Option A: "Evaluate the models by using built-in prompt datasets" is incorrect because built-in datasets may not capture the company's specific stylistic requirements.

? Option C: "Use public model leaderboards to identify the model" is incorrect as leaderboards typically measure model performance on standard benchmarks, not on stylistic preferences.

? Option D: "Use the model InvocationLatency runtime metrics in Amazon CloudWatch" is incorrect because latency metrics do not provide any information about the style of the model's responses.

AWS AI Practitioner References:

? Model Evaluation Techniques on AWS: AWS suggests using human evaluators to assess qualitative aspects of model outputs, such as style and tone, to ensure alignment with organizational preferences

NEW QUESTION 7

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 8

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 9

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 10

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 research company implemented a chatbot by using a foundation model (FM) from Amazon Bedrock. The chatbot searches for answers to questions from a large database of research papers.

After multiple prompt engineering attempts, the company notices that the FM is performing poorly because of the complex scientific terms in the research papers. How can the company improve the performance of the chatbot?

- A. Use few-shot prompting to define how the FM can answer the questions.
- B. Use domain adaptation fine-tuning to adapt the FM to complex scientific terms.
- C. Change the FM inference parameters.
- D. Clean the research paper data to remove complex scientific terms.

Answer: B

Explanation:

Domain adaptation fine-tuning involves training a foundation model (FM) further using a specific dataset that includes domain-specific terminology and content, such as scientific terms in research papers. This process allows the model to better understand and handle complex terminology, improving its performance on specialized tasks.

? Option B (Correct): "Use domain adaptation fine-tuning to adapt the FM to complex scientific terms": This is the correct answer because fine-tuning the model on domain-specific data helps it learn and adapt to the specific language and terms used in the research papers, resulting in better performance.

? Option A: "Use few-shot prompting to define how the FM can answer the questions" is incorrect because while few-shot prompting can help in certain scenarios, it is less effective than fine-tuning for handling complex domain-specific terms.

? Option C: "Change the FM inference parameters" is incorrect because adjusting inference parameters will not resolve the issue of the model's lack of understanding of complex scientific terminology.

? Option D: "Clean the research paper data to remove complex scientific terms" is incorrect because removing the complex terms would result in the loss of important information and context, which is not a viable solution.

AWS AI Practitioner References:

? Domain Adaptation in Amazon Bedrock: AWS recommends fine-tuning models with domain-specific data to improve their performance on specialized tasks involving unique terminology.

NEW QUESTION 13

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 15

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 20

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 23

A company wants to use language models to create an application for inference on edge devices. The inference must have the lowest latency possible. Which solution will meet these requirements?

- A. Deploy optimized small language models (SLMs) on edge devices.
- B. Deploy optimized large language models (LLMs) on edge devices.
- C. Incorporate a centralized small language model (SLM) API for asynchronous communication with edge devices.
- D. Incorporate a centralized large language model (LLM) API for asynchronous communication with edge devices.

Answer: A

Explanation:

To achieve the lowest latency possible for inference on edge devices, deploying optimized small language models (SLMs) is the most effective solution. SLMs require fewer

resources and have faster inference times, making them ideal for deployment on edge devices where processing power and memory are limited.

? Option A (Correct): "Deploy optimized small language models (SLMs) on edge devices": This is the correct answer because SLMs provide fast inference with low latency, which is crucial for edge deployments.

? Option B: "Deploy optimized large language models (LLMs) on edge devices" is incorrect because LLMs are resource-intensive and may not perform well on edge devices due to their size and computational demands.

? Option C: "Incorporate a centralized small language model (SLM) API for asynchronous communication with edge devices" is incorrect because it introduces network latency due to the need for communication with a centralized server.

? Option D: "Incorporate a centralized large language model (LLM) API for asynchronous communication with edge devices" is incorrect for the same reason, with even greater latency due to the larger model size.

AWS AI Practitioner References:

? Optimizing AI Models for Edge Devices on AWS: AWS recommends using small, optimized models for edge deployments to ensure minimal latency and efficient performance.

NEW QUESTION 26

A company has thousands of customer support interactions per day and wants to analyze these interactions to identify frequently asked questions and develop insights.

Which AWS service can the company use to meet this requirement?

- A. Amazon Lex
- B. Amazon Comprehend
- C. Amazon Transcribe
- D. Amazon Translate

Answer: B

Explanation:

Amazon Comprehend is the correct service to analyze customer support interactions and identify frequently asked questions and insights.

? Amazon Comprehend:

? Why Option B is Correct:

? Why Other Options are Incorrect:

NEW QUESTION 31

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 36

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 39

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 43

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.

? 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.

? 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 48

A company wants to deploy a conversational chatbot to answer customer questions. The chatbot is based on a fine-tuned Amazon SageMaker JumpStart model. The application must comply with multiple regulatory frameworks.

Which capabilities can the company show compliance for? (Select TWO.)

- A. Auto scaling inference endpoints
- B. Threat detection

- C. Data protection
- D. Cost optimization
- E. Loosely coupled microservices

Answer: BC

Explanation:

To comply with multiple regulatory frameworks, the company must ensure data protection and threat detection. Data protection involves safeguarding sensitive customer information, while threat detection identifies and mitigates security threats to the application.

? Option C (Correct): "Data protection": This is correct because data protection is critical for compliance with privacy and security regulations.

? Option B (Correct): "Threat detection": This is correct because detecting and mitigating threats is essential to maintaining the security posture required for regulatory compliance.

? Option A: "Auto scaling inference endpoints" is incorrect because auto-scaling does not directly relate to regulatory compliance.

? Option D: "Cost optimization" is incorrect because it is focused on managing expenses, not compliance.

? Option E: "Loosely coupled microservices" is incorrect because this architectural approach does not directly address compliance requirements.

AWS AI Practitioner References:

? AWS Compliance Capabilities: AWS offers services and tools, such as data protection and threat detection, to help companies meet regulatory requirements for security and privacy.

NEW QUESTION 53

A security company is using Amazon Bedrock to run foundation models (FMs). The company wants to ensure that only authorized users invoke the models. The company needs to identify any unauthorized access attempts to set appropriate AWS Identity and Access Management (IAM) policies and roles for future iterations of the FMs.

Which AWS service should the company use to identify unauthorized users that are trying to access Amazon Bedrock?

- A. AWS Audit Manager
- B. AWS CloudTrail
- C. Amazon Fraud Detector
- D. AWS Trusted Advisor

Answer: B

Explanation:

AWS CloudTrail is a service that enables governance, compliance, and operational and risk auditing of your AWS account. It tracks API calls and identifies unauthorized access attempts to AWS resources, including Amazon Bedrock.

? AWS CloudTrail:

? Why Option B is Correct:

? Why Other Options are Incorrect:

Thus, B is the correct answer for identifying unauthorized users attempting to access Amazon Bedrock.

NEW QUESTION 57

A company wants to use generative AI to increase developer productivity and software development. The company wants to use Amazon Q Developer.

What can Amazon Q Developer do to help the company meet these requirements?

- A. Create software snippets, reference tracking, and open-source license tracking.
- B. Run an application without provisioning or managing servers.
- C. Enable voice commands for coding and providing natural language search.
- D. Convert audio files to text documents by using ML models.

Answer: A

Explanation:

Amazon Q Developer is a tool designed to assist developers in increasing productivity by generating code snippets, managing reference tracking, and handling open-source license tracking. These features help developers by automating parts of the software development process.

? Option A (Correct): "Create software snippets, reference tracking, and open-source license tracking": This is the correct answer because these are key features that help developers streamline and automate tasks, thus improving productivity.

? Option B: "Run an application without provisioning or managing servers" is incorrect as it refers to AWS Lambda or AWS Fargate, not Amazon Q Developer.

? Option C: "Enable voice commands for coding and providing natural language search" is incorrect because this is not a function of Amazon Q Developer.

? Option D: "Convert audio files to text documents by using ML models" is incorrect as this refers to Amazon Transcribe, not Amazon Q Developer.

AWS AI Practitioner References:

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

NEW QUESTION 59

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 61

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 66

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 70

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 74

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 79

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 80

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 81

A company is using a pre-trained large language model (LLM) to build a chatbot for product recommendations. The company needs the LLM outputs to be short and written in a specific language.

Which solution will align the LLM response quality with the company's expectations?

- A. Adjust the prompt.
- B. Choose an LLM of a different size.
- C. Increase the temperature.
- D. Increase the Top K value.

Answer: A

Explanation:

Adjusting the prompt is the correct solution to align the LLM outputs with the company's expectations for short, specific language responses.

? Adjust the Prompt:

? Why Option A is Correct:

? Why Other Options are Incorrect:

NEW QUESTION 86

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 89

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 93

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 94

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 95

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 99

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 103

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 108

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 113

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 117

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

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

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

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

Answer: B

Explanation:

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

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

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

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

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

AWS AI Practitioner References:

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

NEW QUESTION 121

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 122

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 123

.....

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](#)