

AIF-C01 Dumps

AWS Certified AI Practitioner

<https://www.certleader.com/AIF-C01-dumps.html>



NEW QUESTION 1

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

Which consideration will inform the company's decision?

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

Answer: B

Explanation:

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

? Context Window:

? Why Option B is Correct:

? Why Other Options are Incorrect:

NEW QUESTION 2

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 company has developed an ML model for image classification. The company wants to deploy the model to production so that a web application can use the model.

The company needs to implement a solution to host the model and serve predictions without managing any of the underlying infrastructure.

Which solution will meet these requirements?

- A. Use Amazon SageMaker Serverless Inference to deploy the model.
- B. Use Amazon CloudFront to deploy the model.
- C. Use Amazon API Gateway to host the model and serve predictions.
- D. Use AWS Batch to host the model and serve predictions.

Answer: A

Explanation:

Amazon SageMaker Serverless Inference is the correct solution for deploying an ML model to production in a way that allows a web application to use the model without the need to manage the underlying infrastructure.

? Amazon SageMaker Serverless Inference provides a fully managed environment

for deploying machine learning models. It automatically provisions, scales, and manages the infrastructure required to host the model, removing the need for the company to manage servers or other underlying infrastructure.

? Why Option A is Correct:

? Why Other Options are Incorrect:

Thus, A is the correct answer, as it aligns with the requirement of deploying an ML model without managing any underlying infrastructure.

NEW QUESTION 4

A company wants to display the total sales for its top-selling products across various retail locations in the past 12 months.

Which AWS solution should the company use to automate the generation of graphs?

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

Answer: C

Explanation:

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

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

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

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

AWS AI Practitioner References:

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

? Business Intelligence (BI) on AWS: AWS services such as Amazon QuickSight

provide business intelligence capabilities, including automated reporting and visualization features, which are ideal for companies seeking to visualize data like sales trends over time.

NEW QUESTION 5

A medical company deployed a disease detection model on Amazon Bedrock. To comply with privacy policies, the company wants to prevent the model from including personal patient information in its responses. The company also wants to receive notification when policy violations occur.

Which solution meets these requirements?

- A. Use Amazon Macie to scan the model's output for sensitive data and set up alerts for potential violations.
- B. Configure AWS CloudTrail to monitor the model's responses and create alerts for any detected personal information.
- C. Use Guardrails for Amazon Bedrock to filter content
- D. Set up Amazon CloudWatch alarms for notification of policy violations.
- E. Implement Amazon SageMaker Model Monitor to detect data drift and receive alerts when model quality degrades.

Answer: C

Explanation:

Guardrails for Amazon Bedrock provide mechanisms to filter and control the content generated by models to comply with privacy and policy requirements. Using guardrails ensures that sensitive or personal information is not included in the model's responses. Additionally, integrating Amazon CloudWatch alarms allows for real-time notification when a policy violation occurs.

? Option C (Correct): "Use Guardrails for Amazon Bedrock to filter content. Set up

Amazon CloudWatch alarms for notification of policy violations": This is the correct answer because it directly addresses both the prevention of policy violations and the requirement to receive notifications when such violations occur.

? Option A: "Use Amazon Macie to scan the model's output for sensitive data" is

incorrect because Amazon Macie is designed to monitor data in S3, not to filter real-time model outputs.

? Option B: "Configure AWS CloudTrail to monitor the model's responses" is

incorrect because CloudTrail tracks API activity and is not suited for content moderation.

? Option D: "Implement Amazon SageMaker Model Monitor to detect data drift" is

incorrect because data drift detection does not address content moderation or privacy compliance.

AWS AI Practitioner References:

? Guardrails in Amazon Bedrock: AWS provides guardrails to ensure AI models comply with content policies, and using CloudWatch for alerting integrates monitoring capabilities.

NEW QUESTION 6

An AI practitioner wants to use a foundation model (FM) to design a search application. The search application must handle queries that have text and images.

Which type of FM should the AI practitioner use to power the search application?

- A. Multi-modal embedding model
- B. Text embedding model
- C. Multi-modal generation model
- D. Image generation model

Answer: A

Explanation:

A multi-modal embedding model is the correct type of foundation model (FM) for powering a search application that handles queries containing both text and images.

? Multi-Modal Embedding Model:

? Why Option A is Correct:

? Why Other Options are Incorrect:

NEW QUESTION 7

A company 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 8

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 the harmonic mean of precision and recall, making it a balanced metric for evaluating model performance when there is an imbalance between false positives and false negatives. Speed, cost, and energy efficiency are unrelated to the F1 score. References: AWS Foundation Models Guide.

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

A company wants to assess the costs that are associated with using a large language model (LLM) to generate inferences. The company wants to use Amazon Bedrock to build generative AI applications.

Which factor will drive the inference costs?

- A. Number of tokens consumed
- B. Temperature value
- C. Amount of data used to train the LLM
- D. Total training time

Answer: A

Explanation:

In generative AI models, such as those built on Amazon Bedrock, inference costs are driven by the number of tokens processed. A token can be as short as one character or as long as one word, and the more tokens consumed during the inference process, the higher the cost.

? Option A (Correct): "Number of tokens consumed": This is the correct answer

because the inference cost is directly related to the number of tokens processed by the model.

? Option B: "Temperature value" is incorrect as it affects the randomness of the

model's output but not the cost directly.

? Option C: "Amount of data used to train the LLM" is incorrect because training data size affects training costs, not inference costs.

? Option D: "Total training time" is incorrect because it relates to the cost of training the model, not the cost of inference.

AWS AI Practitioner References:

? Understanding Inference Costs on AWS: AWS documentation highlights that inference costs for generative models are largely based on the number of tokens processed.

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 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 16

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

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

- A. Using a third-party enterprise application that has embedded generative AI features.
- B. Building an application by using an existing third-party generative AI foundation model (FM).
- C. Refining an existing third-party generative AI foundation model (FM) by fine-tuning the model by using data specific to the business.
- D. Building and training a generative AI model from scratch by using specific data that a customer owns.

Answer: D

Explanation:

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

? Option D (Correct): "Building and training a generative AI model from scratch by using specific data that a customer owns": This is the correct answer because it involves complete ownership of the model, data, and infrastructure, giving the company the highest level of responsibility for security.

? Option A: "Using a third-party enterprise application that has embedded generative AI features" is incorrect as the company has minimal control over the security of the AI features embedded within a third-party application.

? Option B: "Building an application using an existing third-party generative AI foundation model (FM)" is incorrect because security responsibilities are shared with the third-party model provider.

? Option C: "Refining an existing third-party generative AI FM by fine-tuning the model with business-specific data" is incorrect as the foundation model and part of the security responsibilities are still managed by the third party.

AWS AI Practitioner References:

? Generative AI Security Scoping Matrix on AWS: AWS provides a security responsibility matrix that outlines varying levels of control and responsibility depending on the approach to developing and using AI models.

NEW QUESTION 20

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 24

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 27

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 and use an AI model responsibly, especially in sensitive applications like loan approvals, it's crucial to address potential biases and ensure transparency:

? Detect imbalances or disparities in the data (Option A): Analyzing the training data

for imbalances or disparities is essential. Imbalanced data can lead to models that are biased towards the majority class, potentially disadvantaging certain groups.

By identifying and mitigating these imbalances, the company can reduce the risk of biased predictions.

? Evaluate the model's behavior to provide transparency to stakeholders (Option C):

Regularly assessing the model's outputs and decision-making processes allows the company to understand how decisions are made. This evaluation fosters transparency, enabling the company to explain model behavior to stakeholders

and ensure that the model operates as intended without unintended biases. Options B, D, and E, while relevant to model performance and evaluation, do not directly address the responsible use of AI concerning bias and transparency.

Reference: AWS Certified AI Practitioner Exam Guide

NEW QUESTION 28

A company is building a customer service chatbot. The company wants the chatbot to improve its responses by learning from past interactions and online resources.

Which AI learning strategy provides this self-improvement capability?

- A. Supervised learning with a manually curated dataset of good responses and bad responses
- B. Reinforcement learning with rewards for positive customer feedback
- C. Unsupervised learning to find clusters of similar customer inquiries
- D. Supervised learning with a continuously updated FAQ database

Answer: B

Explanation:

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

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

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

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

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

AWS AI Practitioner References:

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

NEW QUESTION 31

A company has built an image classification model to predict plant diseases from photos of plant leaves. The company wants to evaluate how many images the model classified correctly.

Which evaluation metric should the company use to measure the model's performance?

- A. R-squared score
- B. Accuracy
- C. Root mean squared error (RMSE)
- D. Learning rate

Answer: B

Explanation:

Accuracy is the most appropriate metric to measure the performance of an image classification model. It indicates the percentage of correctly classified images out of the total number of images. In the context of classifying plant diseases from images, accuracy will help the company determine how well the model is performing by showing how many images were correctly classified.

? Option B (Correct): "Accuracy": This is the correct answer because accuracy measures the proportion of correct predictions made by the model, which is suitable for evaluating the performance of a classification model.

? Option A: "R-squared score" is incorrect as it is used for regression analysis, not classification tasks.

? Option C: "Root mean squared error (RMSE)" is incorrect because it is also used for regression tasks to measure prediction errors, not for classification accuracy.

? Option D: "Learning rate" is incorrect as it is a hyperparameter for training, not a performance metric.

AWS AI Practitioner References:

? Evaluating Machine Learning Models on AWS: AWS documentation emphasizes the use of appropriate metrics, like accuracy, for classification tasks.

NEW QUESTION 33

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 38

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 42

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 47

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

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

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

Answer: A

Explanation:

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

Explanation of Each Option:

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

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

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

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

AWS AI Practitioner References:

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

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

NEW QUESTION 50

A company wants to use a large language model (LLM) on Amazon Bedrock for sentiment analysis. The company needs the LLM to produce more consistent responses to the same input prompt.

Which adjustment to an inference parameter should the company make to meet these requirements?

- A. Decrease the temperature value
- B. Increase the temperature value
- C. Decrease the length of output tokens
- D. Increase the maximum generation length

Answer: A

Explanation:

The temperature parameter in a large language model (LLM) controls the randomness of the model's output. A lower temperature value makes the output more deterministic and consistent, meaning that the model is less likely to produce different results for the same input prompt.

? Option A (Correct): "Decrease the temperature value": This is the correct answer

because lowering the temperature reduces the randomness of the responses, leading to more consistent outputs for the same input.

? Option B: "Increase the temperature value" is incorrect because it would make the output more random and less consistent.

? Option C: "Decrease the length of output tokens" is incorrect as it does not directly affect the consistency of the responses.

? Option D: "Increase the maximum generation length" is incorrect because this adjustment affects the output length, not the consistency of the model's responses.

AWS AI Practitioner References:

? Understanding Temperature in Generative AI Models: AWS documentation explains that adjusting the temperature parameter affects the model's output randomness, with lower values providing more consistent outputs.

NEW QUESTION 51

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 55

An AI practitioner has a database of animal photos. The AI practitioner wants to automatically identify and categorize the animals in the photos without manual human effort.

Which strategy meets these requirements?

- A. Object detection
- B. Anomaly detection
- C. Named entity recognition
- D. Inpainting

Answer: A

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 56

A company is implementing the Amazon Titan foundation model (FM) by using Amazon Bedrock. The company needs to supplement the model by using relevant data from the company's private data sources.

Which solution will meet this requirement?

- A. Use a different FM
- B. Choose a lower temperature value
- C. Create an Amazon Bedrock knowledge base
- D. Enable model invocation logging

Answer: C

Explanation:

Creating an Amazon Bedrock knowledge base allows the integration of external or private data sources with a foundation model (FM) like Amazon Titan. This integration helps supplement the model with relevant data from the company's private data sources to enhance its responses.

? Option C (Correct): "Create an Amazon Bedrock knowledge base": This is the correct answer as it enables the company to incorporate private data into the FM to improve its effectiveness.

? Option A: "Use a different FM" is incorrect because it does not address the need to supplement the current model with private data.

? Option B: "Choose a lower temperature value" is incorrect as it affects output randomness, not the integration of private data.

? Option D: "Enable model invocation logging" is incorrect because logging does not help in supplementing the model with additional data.

AWS AI Practitioner References:

? Amazon Bedrock and Knowledge Integration: AWS explains how creating a knowledge base allows Amazon Bedrock to use external data sources to improve the FM??s relevance and accuracy.

NEW QUESTION 59

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 Canva
- 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 64

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

- A. Supervised learning
- B. Unsupervised learning
- C. Reinforcement learning
- D. Reinforcement learning from human feedback (RLHF)

Answer: B

Explanation:

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

? Unsupervised Learning:

? Why Option B is Correct:

? Why Other Options are Incorrect:

NEW QUESTION 66

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 71

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

- A. Data leakage
- B. Hallucination
- C. Overfitting
- D. Underfitting

Answer: B

Explanation:

In the context of AI, "hallucination" refers to the phenomenon where a model generates outputs that are plausible-sounding but are not grounded in reality or the training data. This problem often occurs with large language models (LLMs) when they create information that sounds correct but is actually incorrect or fabricated.

? Option B (Correct): "Hallucination": This is the correct answer because the

problem described involves generating content that sounds factual but is incorrect, which is characteristic of hallucination in generative AI models.

? Option A: "Data leakage" is incorrect as it involves the model accidentally learning

from data it shouldn't have access to, which does not match the problem of generating incorrect content.

? Option C: "Overfitting" is incorrect because overfitting refers to a model that has learned the training data too well, including noise, and performs poorly on new data.

? Option D: "Underfitting" is incorrect because underfitting occurs when a model is too simple to capture the underlying patterns in the data, which is not the issue here.

AWS AI Practitioner References:

? Large Language Models on AWS: AWS discusses the challenge of hallucination in large language models and emphasizes techniques to mitigate it, such as using guardrails and fine-tuning.

NEW QUESTION 73

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 75

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 80

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 85

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 86

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 88

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 90

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

- A. Supervised learning.
- B. Unsupervised learning.

- C. Reinforcement learning.
- D. Active learning.

Answer: A

Explanation:

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

NEW QUESTION 91

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 95

An AI practitioner is building a model to generate images of humans in various professions. The AI practitioner discovered that the input data is biased and that specific attributes affect the image generation and create bias in the model. Which technique will solve the problem?

- A. Data augmentation for imbalanced classes
- B. Model monitoring for class distribution
- C. Retrieval Augmented Generation (RAG)
- D. Watermark detection for images

Answer: A

Explanation:

Data augmentation for imbalanced classes is the correct technique to address bias in input data affecting image generation.

? Data Augmentation for Imbalanced Classes:

? Why Option A is Correct:

? Why Other Options are Incorrect:

NEW QUESTION 100

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 102

Which option is a benefit of ongoing pre-training when fine-tuning a foundation model (FM)?

- A. Helps decrease the model's complexity
- B. Improves model performance over time
- C. Decreases the training time requirement
- D. Optimizes model inference time

Answer: B

Explanation:

Ongoing pre-training when fine-tuning a foundation model (FM) improves model performance over time by continuously learning from new data.

? Ongoing Pre-Training:

? Why Option B is Correct:

? Why Other Options are Incorrect:

NEW QUESTION 103

.....

Thank You for Trying Our Product

* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

* One year free update

You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

We currently serve more than 30,000,000 customers.

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your AIF-C01 Exam with Our Prep Materials Via below:

<https://www.certleader.com/AIF-C01-dumps.html>