Amazon-Web-Services

Exam Questions AIF-C01

AWS Certified AI Practitioner

NEW QUESTION 1
A company wants to use a large language model (LLM) on Amazon Bedrock for sentiment analysis. The company wants to know how much information can fit into one prompt.
Which consideration will inform the company's decision?

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

Answer: B

Explanation:
The context window determines how much information can fit into a single prompt when using a large language model (LLM) like those on Amazon Bedrock.
? Context Window:
? Why Option B is Correct:
? Why Other Options are Incorrect:

NEW QUESTION 2
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 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 4
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 5
How can companies use large language models (LLMs) securely on Amazon Bedrock?

A. Design clear and specific prompt
B. Configure AWS Identity and Access Management (IAM) roles and policies by using least privilege access.
C. Enable AWS Audit Manager for automatic model evaluation jobs.
D. Enable Amazon Bedrock automatic model evaluation jobs.
E. Use Amazon CloudWatch Logs to make models explainable and to monitor for bias.

Answer: A

Explanation:
To securely use large language models (LLMs) on Amazon Bedrock, companies should design clear and specific prompts to avoid unintended outputs and ensure proper configuration of AWS Identity and Access Management (IAM) roles and policies with the principle of least privilege. This approach limits access to sensitive resources and minimizes the potential impact of security incidents.
? Option A (Correct): "Design clear and specific prompts. Configure AWS Identity
and Access Management (IAM) roles and policies by using least privilege access": This is the correct answer as it directly addresses both security practices in prompt design and access management.
? Option B: "Enable AWS Audit Manager for automatic model evaluation jobs" is
incorrect because Audit Manager is for compliance and auditing, not directly related to secure LLM usage.
? Option C: "Enable Amazon Bedrock automatic model evaluation jobs" is incorrect
because Bedrock does not provide automatic model evaluation jobs specifically for security purposes.
? Option D: "Use Amazon CloudWatch Logs to make models explainable and to
monitor for bias" is incorrect because CloudWatch Logs are used for monitoring and not directly for making models explainable or secure.
AWS AI Practitioner References:
? Secure AI Practices on AWS: AWS recommends configuring IAM roles and using least privilege access to ensure secure usage of AI models.

NEW QUESTION 6
A company wants to build an ML model by using Amazon SageMaker. The company needs to share and manage variables for model development across multiple teams.
Which SageMaker feature meets these requirements?

A. Amazon SageMaker Feature Store
B. Amazon SageMaker Data Wrangler
C. Amazon SageMaker Clarify
D. Amazon SageMaker Model Cards

Answer: A

Explanation:
Amazon SageMaker Feature Store is the correct solution for sharing and managing variables (features) across multiple teams during model development.
? Amazon SageMaker Feature Store:
? Why Option A is Correct:
? Why Other Options are Incorrect:

NEW QUESTION 7
A company wants to classify human genes into 20 categories based on gene characteristics. The company needs an ML algorithm to document how the inner mechanism of the model affects the output.
Which ML algorithm meets these requirements?

A. Decision trees
B. Linear regression
C. Logistic regression
D. Neural networks

Answer: A

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

NEW QUESTION 8
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 9
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
A company wants to develop a large language model (LLM) application by using Amazon Bedrock and customer data that is uploaded to Amazon S3. The company's security policy states that each team can access data for only the team's own customers.
Which solution will meet these requirements?

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

Answer: A

Explanation:
To comply with the company's security policy, which restricts each team to access data for only their own customers, creating an Amazon Bedrock custom service role for each team is the correct solution.
? Custom Service Role Per Team:
? Why Option A is Correct:
? Why Other Options are Incorrect:
Thus, A is the correct answer to meet the company's security requirements.

NEW QUESTION 10
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 12
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 16
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 21
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 25
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 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 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 33
A company has installed a security camera. The company uses an ML model to evaluate the security camera footage for potential thefts. The company has discovered that the model disproportionately flags people who are members of a specific ethnic group.
Which type of bias is affecting the model output?

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

Answer: B

Explanation:
Sampling bias is the correct type of bias affecting the model output when it disproportionately flags people from a specific ethnic group.
? Sampling Bias:
? Why Option B is Correct:
? Why Other Options are Incorrect:

NEW QUESTION 36
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 41
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 44
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 47
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 52
A company is building an ML model. The company collected new data and analyzed the data by creating a correlation matrix, calculating statistics, and visualizing the data.
Which stage of the ML pipeline is the company currently in?

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

Answer: C

Explanation:
Exploratory data analysis (EDA) involves understanding the data by visualizing it, calculating statistics, and creating correlation matrices. This stage helps identify patterns, relationships, and anomalies in the data, which can guide further steps in the ML pipeline.
? Option C (Correct): "Exploratory data analysis": This is the correct answer as the
tasks described (correlation matrix, calculating statistics, visualizing data) are all part of the EDA process.
? Option A: "Data pre-processing" is incorrect because it involves cleaning and
transforming data, not initial analysis.
? Option B: "Feature engineering" is incorrect because it involves creating new features from raw data, not analyzing the data's existing structure.
? Option D: "Hyperparameter tuning" is incorrect because it refers to optimizing model parameters, not analyzing the data.
AWS AI Practitioner References:
? Stages of the Machine Learning Pipeline: AWS outlines EDA as the initial phase of understanding and exploring data before moving to more specific preprocessing, feature engineering, and model training stages.

NEW QUESTION 54
A company uses Amazon SageMaker for its ML pipeline in a production environment. The company has large input data sizes up to 1 GB and processing times up to 1 hour. The company needs near real-time latency.
Which SageMaker inference option meets these requirements?

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

Answer: A

Explanation:
Real-time inference is designed to provide immediate, low-latency predictions, which is necessary when the company requires near real-time latency for its ML models. This option is optimal when there is a need for fast responses, even with large input data sizes and substantial processing times.
? Option A (Correct): "Real-time inference": This is the correct answer because it
supports low-latency requirements, which are essential for real-time applications where quick response times are needed.
? Option B: "Serverless inference" is incorrect because it is more suited for intermittent, small-scale inference workloads, not for continuous, large-scale, low-latency needs.
? Option C: "Asynchronous inference" is incorrect because it is used for workloads that do not require immediate responses.
? Option D: "Batch transform" is incorrect as it is intended for offline, large-batch processing where immediate response is not necessary.
AWS AI Practitioner References:
? Amazon SageMaker Inference Options: AWS documentation describes real-time inference as the best solution for applications that require immediate prediction results with low latency.

NEW QUESTION 56
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 61
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 66
A company wants to develop an educational game where users answer questions such as the following: "A jar contains six red, four green, and three yellow marbles. What is the probability of choosing a green marble from the jar?"
Which solution meets these requirements with the LEAST operational overhead?

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

Answer: C

Explanation:
The problem involves a simple probability calculation that can be handled efficiently by straightforward mathematical rules and computations. Using machine learning techniques would introduce unnecessary complexity and operational overhead.
? Option C (Correct): "Use code that will calculate probability by using simple rules and computations": This is the correct answer because it directly solves the problem with minimal overhead, using basic probability rules.
? Option A: "Use supervised learning to create a regression model" is incorrect as it overcomplicates the solution for a simple probability problem.
? Option B: "Use reinforcement learning to train a model" is incorrect because reinforcement learning is not needed for a simple probability calculation.
? Option D: "Use unsupervised learning to create a model" is incorrect as unsupervised learning is not applicable to this task.
AWS AI Practitioner References:
? Choosing the Right Solution for AI Tasks: AWS recommends using the simplest and most efficient approach to solve a given problem, avoiding unnecessary machine learning techniques for straightforward tasks.

NEW QUESTION 71
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 74
A company is using few-shot prompting on a base model that is hosted on Amazon Bedrock. The model currently uses 10 examples in the prompt. The model is invoked once daily and is performing well. The company wants to lower the monthly cost.
Which solution will meet these requirements?

A. Customize the model by using fine-tuning.
B. Decrease the number of tokens in the prompt.
C. Increase the number of tokens in the prompt.
D. Use Provisioned Throughput.

Answer: B

Explanation:
Decreasing the number of tokens in the prompt reduces the cost associated with using an LLM model on Amazon Bedrock, as costs are often based on the number of tokens processed by the model.
? Token Reduction Strategy:
? Why Option B is Correct:
? Why Other Options are Incorrect:

NEW QUESTION 78
A loan company is building a generative AI-based solution to offer new applicants discounts based on specific business criteria. The company wants to build and use an AI model responsibly to minimize bias that could negatively affect some customers.
Which actions should the company take to meet these requirements? (Select TWO.)

A. Detect imbalances or disparities in the data.
B. Ensure that the model runs frequently.
C. Evaluate the model's behavior so that the company can provide transparency to stakeholders.
D. Use the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) technique to ensure that the model is 100% accurate.
E. Ensure that the model's inference time is within the accepted limits.

Answer: AC

Explanation:
To build an AI model responsibly and minimize bias, it is essential to ensure fairness and transparency throughout the model development and deployment process. This involves detecting and mitigating data imbalances and thoroughly evaluating the model's behavior to understand its impact on different groups.
? Option A (Correct): "Detect imbalances or disparities in the data": This is correct because identifying and addressing data imbalances or disparities is a critical step in reducing bias. AWS provides tools like Amazon SageMaker Clarify to detect bias during data preprocessing and model training.
? Option C (Correct): "Evaluate the model's behavior so that the company can provide transparency to stakeholders": This is correct because evaluating the model's behavior for fairness and accuracy is key to ensuring that stakeholders understand how the model makes decisions. Transparency is a crucial aspect of responsible AI.
? Option B: "Ensure that the model runs frequently" is incorrect because the frequency of model runs does not address bias.
? Option D: "Use the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) technique to ensure that the model is 100% accurate" is incorrect because ROUGE is a metric for evaluating the quality of text summarization models, not for minimizing bias.
? Option E: "Ensure that the model's inference time is within the accepted limits" is incorrect as it relates to performance, not bias reduction.
AWS AI Practitioner References:
? Amazon SageMaker Clarify: AWS offers tools such as SageMaker Clarify for detecting bias in datasets and models, and for understanding model behavior to ensure fairness and transparency.
? Responsible AI Practices: AWS promotes responsible AI by advocating for fairness, transparency, and inclusivity in model development and deployment.

NEW QUESTION 83
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 86
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 91
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 96
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 withthe 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 100
.....

Thank You for Trying Our Product

We offer two products:

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

2nd - Questons 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