Vendor: Amazon

Exam Code:AIF-C01

Exam Name: Amazon AWS Certified AI Practitioner

Version: Demo

QUESTION 1

A company wants to use a large language model (LLM) on Amazon Bedrock for sentiment analysis. The company wants to classify the sentiment of text passages as positive or negative.

Which prompt engineering strategy meets these requirements?

A. Provide examples of text passages with corresponding positive or negative labels in the prompt followed by the new text passage to be classified.

B. Provide a detailed explanation of sentiment analysis and how LLMs work in the prompt.

C. Provide the new text passage to be classified without any additional context or examples.

D. Provide the new text passage with a few examples of unrelated tasks, such as text summarization or question answering.

Correct Answer: A

Providing examples of text passages with corresponding positive or negative labels in the prompt followed by the new text passage to be classified is the correct prompt engineering strategy for using a large language model (LLM) on Amazon Bedrock for sentiment analysis.

QUESTION 2

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.

Correct Answer: A

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.

QUESTION 3

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

Correct Answer: B

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.

QUESTION 4

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

Correct Answer: A

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.

QUESTION 5

A company is building an application that needs to generate synthetic data that is based on existing data.

Which type of model can the company use to meet this requirement?

A. Generative adversarial network (GAN)

B. XGBoost

C. Residual neural network

D. WaveNet

Correct Answer: A

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

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

B:"XGBoost" is a gradient boosting algorithm for classification and regression tasks, not for generating synthetic data. Option C:"Residual neural network" is primarily used for improving the performance of deep networks, not for generating

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

AWS AI Practitioner References:

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

QUESTION 6

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.

Correct Answer: B

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

QUESTION 7

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

Correct Answer: A

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.

QUESTION 8

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

Which AWS service meets these requirements?

A. Amazon S3

B. Amazon Elastic Block Store (Amazon EBS)

C. Amazon Elastic File System (Amazon EFS)

D. AWS Snowcone

Correct Answer: A

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

validation purposes. Option A (Correct): "Amazon S3":This is the correct answer because Amazon S3 is widely used for storing large datasets that are accessed by machine learning models, including those in Amazon Bedrock.

Option B:"Amazon Elastic Block Store (Amazon EBS)" is incorrect because EBS is a block storage service for use with Amazon EC2, not for directly storing datasets for Amazon Bedrock.

Option C:"Amazon Elastic File System (Amazon EFS)" is incorrect as it is primarily used for file storage with shared access by multiple instances. Option D:"AWS Snowcone" is incorrect because it is a physical device for offline data transfer,

not suitable for directly providing data to Amazon Bedrock.

AWS AI Practitioner

References:

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

QUESTION 9

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

Which Amazon SageMaker inference option will meet these requirements?

A. Batch transform

B. Real-time inference

C. Serverless inference

D. Asynchronous inference

Correct Answer: A

Batch transform in Amazon SageMaker is designed for offline processing of large datasets. It is ideal for scenarios where immediate predictions are not required, and the inference can be done on large datasets that are multiple gigabytes in

size. This method processes data in batches, making it suitable for analyzing archived data without the need for real-time access to predictions.

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

Option B:"Real-time inference" is incorrect because it is used for low-latency, real- time prediction needs, which is not required in this case. Option C:"Serverless inference" is incorrect because it is designed for small-scale, intermittent

inference requests, not for large batch processing. Option D:"Asynchronous inference" is incorrect because it is used when immediate predictions are required, but with high throughput, whereas batch transform is more suitable for very large

datasets.

AWS AI Practitioner

References:

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

QUESTION 10

A company is using Amazon SageMaker Studio notebooks to build and train ML models. The company stores the data in an Amazon S3 bucket. The company needs to manage the flow of data from Amazon S3 to SageMaker Studio notebooks.

Which solution will meet this requirement?

A. Use Amazon Inspector to monitor SageMaker Studio.

B. Use Amazon Macie to monitor SageMaker Studio.

C. Configure SageMaker to use a VPC with an S3 endpoint.

D. Configure SageMaker to use S3 Glacier Deep Archive.

Correct Answer: C

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

QUESTION 11

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 Wrangler. Create ML models and demand forecast predictions by using SageMaker built-in algorithms.

C. Import the data into Amazon SageMaker Data Wrangler. Build ML models and demand forecast predictions by using an Amazon Personalize Trending-Now recipe.

D. Import the data into Amazon SageMaker Canvas. Build ML models and demand forecast predictions by selecting the values in the data from SageMaker Canvas.

Correct Answer: D

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.

QUESTION 12

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.

Correct Answer: A

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.