

Q201. An education company waftion. The application will give users the ability to enter text or provide a picture of a question. The application will respond with a written answer and an explanation of the written answer. Which model type meets these requirements?

- A. Computer vision model
- B. Large multi-modal language model
- C. Diffusion model
- D. Text-to-speech model

Answer: B

Q202. In which stage of the generative AI model lifecycle are tests performed to examine the model's accuracy?

- A. Deployment
- B. Data selection
- C. Fine-tuning
- D. Evaluation

Answer: D

Q203. Which statement correctly describes embeddings in generative AI?

- A. Embeddings represent data as high-dimensional vectors that capture semantic relationships.
- B. Embeddings is a technique that searches data to find the most helpful information to answer natural language questions.
- C. Embeddings reduce the hardware requirements of a model by using a less precise data type for the weights and activations.
- D. Embeddings provide the ability to store and retrieve data for generative AI applications.

Answer: A

Q204. A company wants to add generative AI functionality to its application by integrating a large language model (LLM). The responses from the LLM must be as deterministic and as stable as possible. Which solution meets these requirements?

- A. Configure the application to automatically set the temperature parameter to 0 when submitting the prompt to the LLM.
- B. Configure the application to automatically add "make your response deterministic" at the end of the prompt before submitting the prompt to the LLM.
- C. Configure the application to automatically add "make your response deterministic" at the beginning of the prompt before submitting the prompt to the LLM.
- D. Configure the application to automatically set the temperature parameter to 1 when submitting the prompt to the LLM.

Answer: A

Q205. A company needs to select a generative AI model to build an application. The application must provide responses to users in real time. Which model characteristic should the company consider to meet these requirements?

- A. Model complexity
- B. Innovation speed
- C. Inference speed
- D. Training time

Answer: C

Q206. Which term refers to the instructions given to foundation models (FMs) so that the FMs provide a more accurate response to a question?

- A. Prompt
- B. Direction
- C. Dialog
- D. Translation

Answer: A

Q207. A retail company wants to build an ML model to recommend products to customers. The company wants to build the model based on responsible practices. Which practice should the company apply when collecting data to decrease model bias?

- A. Use data from only customers who match the demographics of the company's overall customer base.
- B. Collect data from customers who have a past purchase history.
- C. Ensure that the data is balanced and collected from a diverse group.
- D. Ensure that the data is from a publicly available dataset.

Answer: C

Q208. A company is developing an ML model to predict customer churn. Which evaluation metric will assess the model's performance on a binary classification task such as predicting churn?

- A. F1 score
- B. Mean squared error (MSE)
- C. R-squared
- D. Time used to train the model

Answer: A

Q209. An AI practitioner is evaluating the performance of an Amazon SageMaker model. The AI practitioner must choose a performance metric. The metric must show the ratio of the number of correctly classified items to the total number of correctly and incorrectly classified items. Which metric meets these requirements?

- A. Accuracy
- B. Precision
- C. F1 score
- D. Recall

Answer: A

Q210. An ecommerce company receives multiple gigabytes of customer data daily. The company uses the data to train an ML model to forecast future product demand. The company needs a solution to perform inferences once each day. Which inference type meets these requirements?

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

Answer: A

Q211. A company has developed a generative AI model for customer segmentation. The model has been deployed in the company's production environment for a long time. The company recently noticed some inconsistency in the model's responses. The company wants to evaluate model bias and drift. Which AWS service or feature meets these requirements?

- A. Amazon SageMaker Model Monitor
- B. Amazon SageMaker Clarify
- C. Amazon SageMaker Model Cards
- D. Amazon SageMaker Feature Store

Answer:  B

Q212. A company has signed up for Amazon Bedrock access to build applications. The company wants to restrict employee access to specific models available on Amazon Bedrock. Which solution meets these requirements?

- A. Use AWS Identity and Access Management (IAM) policies to restrict model access.
- B. Use AWS Security Token Service (AWS STS) to generate temporary credentials for model use.
- C. Use AWS Identity and Access Management (IAM) service roles to restrict model subscription.
- D. Use Amazon Inspector to monitor model access.

Answer: A

Q213. Which ML technique uses training data that is labeled with the correct output values?

- A. Supervised learning
- B. Unsupervised learning
- C. Reinforcement learning
- D. Transfer learning

Answer: A

Q214. Which large language model (LLM) parameter controls the number of possible next words or tokens considered at each step of the text generation process?

- A. Maximum tokens
- B. Top K
- C. Temperature
- D. Batch size

Answer: B

Q215. A company is making a chatbot. The chatbot uses Amazon Lex and Amazon OpenSearch Service. The chatbot uses the company's private data to answer questions. The company needs to convert the data into a vector representation before storing the data in a database. Which type of foundation model (FM) meets these requirements?

- A. Text completion model
- B. Instruction following model
- C. Text embeddings model
- D. Image generation model

Answer: C

Q216. A company wants to use a large language model (LLM) to generate product descriptions. The company wants to give the model example descriptions that follow a format. Which prompt engineering technique will generate descriptions that match the format?

- A. Zero-shot prompting
- B. Chain-of-thought prompting
- C. One-shot prompting
- D. Few-shot prompting

Answer: D

Q217. A bank is fine-tuning a large language model (LLM) on Amazon Bedrock to assist customers with questions about their loans. The bank wants to ensure that the model does not reveal any private customer data. Which solution meets these requirements?

- A. Use Amazon Bedrock Guardrails.
- B. Remove personally identifiable information (PII) from the customer data before fine-tuning the LLM.
- C. Increase the Top-K parameter of the LLM.
- D. Store customer data in Amazon S3. Encrypt the data before fine-tuning the LLM.

Answer: B

Q218. A grocery store wants to create a chatbot to help customers find products in the store. The chatbot must check the inventory in real time and provide the product location in the store. Which prompt engineering technique should the store use to build the chatbot?

- A. Zero-shot prompting
- B. Few-shot prompting
- C. Least-to-most prompting
- D. Reasoning and acting (ReAct) prompting

Answer: D

Q219. A company uses a third-party model on Amazon Bedrock to analyze confidential documents. The company is concerned about data privacy. Which statement describes how Amazon Bedrock protects data privacy?

- A. User inputs and model outputs are anonymized and shared with third-party model providers.
- B. User inputs and model outputs are not shared with any third-party model providers.
- C. User inputs are kept confidential, but model outputs are shared with third-party model providers.
- D. User inputs and model outputs are redacted before the inputs and outputs are shared with third-party model providers.

Answer: B

Q220. An animation company wants to provide subtitles for its content. Which AWS service meets this requirement?

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

Answer: C

Q221. An ecommerce company wants to group customers based on their purchase history and preferences to personalize the user experience of the company's application. Which ML technique should the company use?

- A. Classification
- B. Clustering
- C. Regression
- D. Content generation

Answer: B

Q222. A company wants to control employee access to publicly available foundation models (FMs). Which solution meets these requirements?

- A. Analyze cost and usage reports in AWS Cost Explorer.
- B. Download AWS security and compliance documents from AWS Artifact.
- C. Configure Amazon SageMaker JumpStart to restrict discoverable FMs.
- D. Build a hybrid search solution by using Amazon OpenSearch Service.

Answer: C

Q223. A company has set up a translation tool to help its customer service team handle issues from customers around the world. The company wants to evaluate the performance of the translation tool. The company sets up a parallel data process that compares the responses from the tool to responses from actual humans. Both sets of responses are generated on the same set of documents. Which strategy should the company use to evaluate the translation tool?

- A. Use the Bilingual Evaluation Understudy (BLEU) score to estimate the absolute translation quality of the two methods.
- B. Use the Bilingual Evaluation Understudy (BLEU) score to estimate the relative translation quality of the two methods.
- C. Use the BERTScore to estimate the absolute translation quality of the two methods.
- D. Use the BERTScore to estimate the relative translation quality of the two methods.

Answer: B

Q224. An AI practitioner wants to generate more diverse and more creative outputs from a large language model (LLM). How should the AI practitioner adjust the inference parameter?

- A. Increase the temperature value.
- B. Decrease the Top K value.
- C. Increase the response length.
- D. Decrease the prompt length.

Answer: A

Q225. A company has developed custom computer vision models. The company needs a user-friendly interface for data labeling to minimize model mistakes on new real-world data. Which AWS service, feature, or tool meets these requirements?

- A. Amazon SageMaker Ground Truth
- B. Amazon SageMaker Canvas
- C. Amazon Bedrock playground

D. Amazon Bedrock Agents

Answer: A

Q226. A company is integrating AI into its employee recruitment and hiring solution. The company wants to mitigate bias risks and ensure responsible AI practices while prioritizing equitable hiring decisions. Which core dimensions of responsible AI should the company consider? (Choose two.)

- A. Fairness
- B. Tolerance
- C. Flexibility
- D. Open source
- E. Transparency

Answer: AE

Q227. A financial company has deployed an ML model to predict customer churn. The model has been running in production for 1 week. The company wants to evaluate how accurately the model predicts churn compared to actual customer behavior. Which metric meets these requirements?

- A. Root mean squared error (RMSE)
- B. Return on investment (ROI)
- C. F1 score
- D. Bilingual Evaluation Understudy (BLEU) score

Answer: C

Q228. A company has a generative AI application that uses a pre-trained foundation model (FM) on Amazon Bedrock. The company wants the FM to include more context by using company information. Which solution meets these requirements MOST cost-effectively?

- A. Use Amazon Bedrock Knowledge Bases.
- B. Choose a different FM on Amazon Bedrock.
- C. Use Amazon Bedrock Agents.
- D. Deploy a custom model on Amazon Bedrock.

Answer: A

Q229. A company is using Amazon SageMaker to develop AI models. Select the correct SageMaker feature or resource from the following list for each step in the AI model lifecycle workflow. Each SageMaker feature or resource should be selected one time or not at all.

Answer Area

Managing different versions of the model

SageMaker Clarify
SageMaker Model Registry
SageMaker Serverless Inference

Using the current model to make predictions

SageMaker Clarify
SageMaker Model Registry
SageMaker Serverless Inference

Answer:

Managing different versions of the model

SageMaker Clarify
SageMaker Model Registry
SageMaker Serverless Inference

Using the current model to make predictions

SageMaker Clarify
SageMaker Model Registry
SageMaker Serverless Inference

Q230. A food service company wants to collect a dataset to predict customer food preferences. The company wants to ensure that the food preferences of all demographics are included in the data. Which dataset characteristic does this scenario present?

- A. Accuracy
- B. Diversity
- C. Recency bias
- D. Reliability

Answer: B

Q231. A company wants to create a chatbot that answers questions about human resources policies. The company is using a large language model (LLM) and has a large digital documentation base. Which technique should the company use to optimize the generated responses?

- A. Use Retrieval Augmented Generation (RAG).
- B. Use few-shot prompting.
- C. Set the temperature to 1.
- D. Decrease the token size.

Answer: A

Q232. An education company is building a chatbot whose target audience is teenagers. The company is training a custom large language model (LLM). The company wants the chatbot to

speak in the target audience's language style by using creative spelling and shortened words. Which metric will assess the LLM's performance?

- A. F1 score
- B. BERTScore
- C. Recall-Oriented Understudy for Gisting Evaluation (ROUGE)
- D. Bilingual Evaluation Understudy (BLEU) score

Answer: D

Q233. A customer service team is developing an application to analyze customer feedback and automatically classify the feedback into different categories. The categories include product quality, customer service, and delivery experience. Which A1 concept does this scenario present?

- A. Computer vision
- B. Natural language processing (NLP)
- C. Recommendation systems
- D. Fraud detection

Answer: B