

Q151. A company wants to identify harmful language in the comments section of social media posts by using an ML model. The company will not use labeled data to train the model.

Which strategy should the company use to identify harmful language?

- A. Use Amazon Rekognition moderation.
- B. Use Amazon Comprehend toxicity detection.
- C. Use Amazon SageMaker built-in algorithms to train the model.
- D. Use Amazon Polly to monitor comments.

Answer: B

Q152. A media company wants to analyze viewer behavior and demographics to recommend personalized content. The company wants to deploy a customized ML model in its production environment. The company also wants to observe if the model quality drifts over time. Which AWS service or feature meets these requirements?

- A. Amazon Rekognition
- B. Amazon SageMaker Clarify
- C. Amazon Comprehend
- D. Amazon SageMaker Model Monitor

Answer: D

Q153. A company is deploying AI/ML models by using AWS services. The company wants to offer transparency into the models' decision-making processes and provide explanations for the model outputs.

Which AWS service or feature meets these requirements?

- A. Amazon SageMaker Model Cards
- B. Amazon Rekognition
- C. Amazon Comprehend
- D. Amazon Lex

Answer: A

Q154. A manufacturing company wants to create product descriptions in multiple languages. Which AWS service will automate this task?

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

Answer: A

Q155. A company wants more customized responses to its generative AI models' prompts. Select the correct customization methodology from the following list for each use case. Each use case should be selected one time.

The models must be taught a new domain-specific task

▼

Continued pre-training
Data augmentation
Model fine-tuning

A limited amount of labeled data is available and more data is needed

▼

Continued pre-training
Data augmentation
Model fine-tuning

Only unlabeled data is available

▼

Continued pre-training
Data augmentation
Model fine-tuning

Answer:

The models must be taught a new domain-specific task

▼

Continued pre-training
Data augmentation
Model fine-tuning

A limited amount of labeled data is available and more data is needed

▼

Continued pre-training
Data augmentation
Model fine-tuning

Only unlabeled data is available

▼

Continued pre-training
Data augmentation
Model fine-tuning

Q156. Which AWS feature records details about ML instance data for governance and reporting?

- A. Amazon SageMaker Model Cards
- B. Amazon SageMaker Debugger
- C. Amazon SageMaker Model Monitor
- D. Amazon SageMaker JumpStart

Answer: A

Q157. A financial company is using ML to help with some of the company's tasks. Which option is a use of generative AI models?

- A. Summarizing customer complaints
- B. Classifying customers based on product usage
- C. Segmenting customers based on type of investments
- D. Forecasting revenue for certain products

Answer: A

Q158. A medical company wants to develop an AI application that can access structured patient records, extract relevant information, and generate concise summaries. Which solution will meet these requirements?

- A. Use Amazon Comprehend Medical to extract relevant medical entities and relationships. Apply rule-based logic to structure and format summaries.
- B. Use Amazon Personalize to analyze patient engagement patterns. Integrate the output with a general purpose text summarization tool.
- C. Use Amazon Textract to convert scanned documents into digital text. Design a keyword extraction system to generate summaries.
- D. Implement Amazon Kendra to provide a searchable index for medical records. Use a template-based system to format summaries.

Answer: A

Q159. Which option describes embeddings in the context of AI?

- A. A method for compressing large datasets
- B. An encryption method for securing sensitive data
- C. A method for visualizing high-dimensional data
- D. A numerical method for data representation in a reduced dimensionality space

Answer: D

Q160. A company is building an AI application to summarize books of varying lengths. During testing, the application fails to summarize some books. Why does the application fail to summarize some books?

- A. The temperature is set too high.
- B. The selected model does not support fine-tuning.
- C. The Top P value is too high.
- D. The input tokens exceed the model's context size.

Answer: D

Q161. An airline company wants to build a conversational AI assistant to answer customer questions about flight schedules, booking, and payments. The company wants to use large language models (LLMs) and a knowledge base to create a text-based chatbot interface. Which solution will meet these requirements with the LEAST development effort?

- A. Train models on Amazon SageMaker Autopilot.
- B. Develop a Retrieval Augmented Generation (RAG) agent by using Amazon Bedrock.
- C. Create a Python application by using Amazon Q Developer.
- D. Fine-tune models on Amazon SageMaker Jumpstart.

Answer: B

Q162. What is tokenization used for in natural language processing (NLP)?

- A. To encrypt text data
- B. To compress text files
- C. To break text into smaller units for processing
- D. To translate text between languages

Answer: C

Q163. Which option is a characteristic of transformer-based language models?

- A. Transformer-based language models use convolutional layers to apply filters across an input to capture local patterns through filtered views.
- B. Transformer-based language models can process only text data.
- C. Transformer-based language models use self-attention mechanisms to capture contextual relationships.
- D. Transformer-based language models process data sequences one element at a time in cyclic iterations.

Answer: C

Q164. A financial company is using AI systems to obtain customer credit scores as part of the loan application process. The company wants to expand to a new market in a different geographic area. The company must ensure that it can operate in that geographic area. Which compliance laws should the company review?

- A. Local health data protection laws
- B. Local payment card data protection laws
- C. Local education privacy laws
- D. Local algorithm accountability laws

Answer: D

Q165. A company uses Amazon Bedrock for its generative AI application. The company wants to use Amazon Bedrock Guardrails to detect and filter harmful user inputs and model-generated outputs.

Which content categories can the guardrails filter? (Choose two.)

- A. Hate
- B. Politics
- C. Violence
- D. Gambling
- E. Religion

Answer: AC

Q166. Which scenario describes a potential risk and limitation of prompt engineering in the context of a generative AI model?

- A. Prompt engineering does not ensure that the model always produces consistent and deterministic outputs, eliminating the need for validation.
- B. Prompt engineering could expose the model to vulnerabilities such as prompt injection attacks.
- C. Properly designed prompts reduce but do not eliminate the risk of data poisoning or model hijacking.
- D. Prompt engineering does not ensure that the model will consistently generate highly reliable outputs when working with real-world data.

Answer: B

Q167. A publishing company built a Retrieval Augmented Generation (RAG) based solution to give its users the ability to interact with published content. New content is published daily. The company wants to provide a near real-time experience to users. Which steps in the RAG pipeline should the company implement by using offline batch processing to meet these requirements? (Choose two.)

- A. Generation of content embeddings
- B. Generation of embeddings for user queries
- C. Creation of the search index
- D. Retrieval of relevant content
- E. Response generation for the user

Answer: AC

Q168. Which technique breaks a complex task into smaller subtasks that are sent sequentially to a large language model (LLM)?

- A. One-shot prompting
- B. Prompt chaining
- C. Tree of thoughts
- D. Retrieval Augmented Generation (RAG)

Answer: B

Q169. An AI practitioner needs to improve the accuracy of a natural language generation model. The model uses rapidly changing inventory data. Which technique will improve the model's accuracy?

- A. Transfer learning
- B. Federated learning
- C. Retrieval Augmented Generation (RAG)
- D. One-shot prompting

Answer: C

Q170. A company wants to collaborate with several research institutes to develop an AI model. The company needs standardized documentation of model version tracking and a record of model development.

Which solution meets these requirements?

- A. Track the model changes by using Git.
- B. Track the model changes by using Amazon Fraud Detector.
- C. Track the model changes by using Amazon SageMaker Model Cards.
- D. Track the model changes by using Amazon Comprehend.

Answer: C

Q171. A company that uses multiple ML models wants to identify changes in original model quality so that the company can resolve any issues. Which AWS service or feature meets these requirements?

- A. Amazon SageMaker JumpStart
- B. Amazon SageMaker HyperPod
- C. Amazon SageMaker Data Wrangler
- D. Amazon SageMaker Model Monitor

Answer: D

Q172. What is the purpose of chunking in Retrieval Augmented Generation (RAG)?

- A. To avoid database storage limitations for large text documents by storing parts or chunks of the text
- B. To improve efficiency by avoiding the need to convert large text into vector embeddings
- C. To improve the contextual relevancy of results retrieved from the vector index
- D. To decrease the cost of storage by storing parts or chunks of the text

Answer: C

Q173. A company is developing an editorial assistant application that uses generative AI. During the pilot phase, usage is low and application performance is not a concern. The company cannot predict application usage after the application is fully deployed and wants to minimize application costs.

Which solution will meet these requirements?

- A. Use GPU-powered Amazon EC2 instances.
- B. Use Amazon Bedrock with Provisioned Throughput.
- C. Use Amazon Bedrock with On-Demand Throughput.
- D. Use Amazon SageMaker JumpStart.

Answer: C

Q174. A company deployed a Retrieval Augmented Generation (RAG) application on Amazon Bedrock that gathers financial news to distribute in daily newsletters. Users have recently reported politically influenced ideas in the newsletters. Which Amazon Bedrock guardrail can identify and filter this content?

- A. Word filters
- B. Denied topics
- C. Sensitive information filters
- D. Content filters

Answer: B

Q175. A financial company is developing a fraud detection system that flags potential fraud cases in credit card transactions. Employees will evaluate the flagged fraud cases. The company wants to minimize the amount of time the employees spend reviewing flagged fraud cases that are not actually fraudulent.

Which evaluation metric meets these requirements?

- A. Recall
- B. Accuracy
- C. Precision
- D. Lift chart

Answer: C

Q176. A company designed an AI-powered agent to answer customer inquiries based on product manuals.

Which strategy can improve customer confidence levels in the AI-powered agent's responses?

- A. Writing the confidence level in the response
- B. Including referenced product manual links in the response
- C. Designing an agent avatar that looks like a computer
- D. Training the agent to respond in the company's language style

Answer: B

Q177. A hospital developed an AI system to provide personalized treatment recommendations for patients. The AI system must provide the rationale behind the recommendations and make the insights accessible to doctors and patients.

Which human-centered design principle does this scenario present?

- A. Explainability
- B. Privacy and security
- C. Fairness
- D. Data governance

Answer: A

Q178. Which statement presents an advantage of using Retrieval Augmented Generation (RAG) for natural language processing (NLP) tasks?

- A. RAG can use external knowledge sources to generate more accurate and informative responses.
- B. RAG is designed to improve the speed of language model training.
- C. RAG is primarily used for speech recognition tasks.
- D. RAG is a technique for data augmentation in computer vision tasks.

Answer: A

Q179. A company has created a custom model by fine-tuning an existing large language model (LLM) from Amazon Bedrock. The company wants to deploy the model to production and use the model to handle a steady rate of requests each minute. Which solution meets these requirements MOST cost-effectively?

- A. Deploy the model by using an Amazon EC2 compute optimized instance.
- B. Use the model with on-demand throughput on Amazon Bedrock.
- C. Store the model in Amazon S3 and host the model by using AWS Lambda.
- D. Purchase Provisioned Throughput for the model on Amazon Bedrock.

Answer: D

Q180. Which technique involves training AI models on labeled datasets to adapt the models to specific industry terminology and requirements?

- A. Data augmentation

- B. Fine-tuning
- C. Model quantization
- D. Continuous pre-training

Answer: B

Q181. A company is creating an agent for its application by using Amazon Bedrock Agents. The agent is performing well, but the company wants to improve the agent's accuracy by providing some specific examples. Which solution meets these requirements?

- A. Modify the advanced prompts for the agent to include the examples.
- B. Create a guardrail for the agent that includes the examples.
- C. Use Amazon SageMaker Ground Truth to label the examples.
- D. Run a script in AWS Lambda that adds the examples to the training dataset.

Answer: A

Q182. Which option is a benefit of using infrastructure as code (IaC) in machine learning operations (MLOps)?

- A. IaC eliminates the need for hyperparameter tuning.
- B. IaC always provisions powerful compute instances, contributing to the training of more accurate models.
- C. IaC streamlines the deployment of scalable and consistent ML workloads in cloud environments.
- D. IaC minimizes overall expenses by deploying only low-cost instances.

Answer: C

Q183. A company wants to fine-tune a foundation model (FM) to answer questions for a specific domain. The company wants to use instruction-based fine-tuning. How should the company prepare the training data?

- A. Gather company internal documents and industry-specific materials. Merge the documents and materials into a single file.
- B. Collect external company reviews from various online sources. Manually label each review as either positive or negative.
- C. Create pairs of questions and answers that specifically address topics related to the company's industry domain.
- D. Create few-shot prompts to instruct the model to answer only domain knowledge.

Answer: C

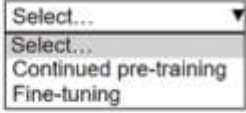
Q184. Which ML technique ensures data compliance and privacy when training AI models on AWS?

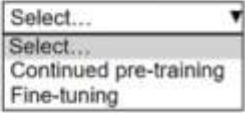
- A. Reinforcement learning
- B. Transfer learning
- C. Federated learning
- D. Unsupervised learning

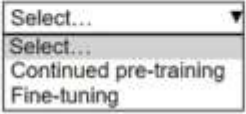
Answer: C

Q185. A company needs to customize a base model that is hosted on Amazon Bedrock.

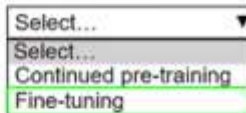
Select the correct model customization method from the following list of company requirements. Each model customization method should be selected one or more times.

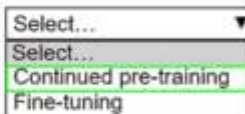
The company wants to improve the model's performance on specific tasks and examples. 

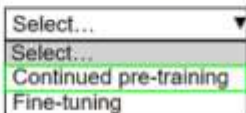
The company wants to improve the model's domain knowledge by providing specific documents. 

The company wants to retrain the model by using more unlabeled data over time. 

Answer:

The company wants to improve the model's performance on specific tasks and examples. 

The company wants to improve the model's domain knowledge by providing specific documents. 

The company wants to retrain the model by using more unlabeled data over time. 

Q186. A manufacturing company has an application that ingests consumer complaints from publicly available sources. The application uses complex hard-coded logic to process the complaints. The company wants to scale this logic across markets and product lines.

Which advantage do generative AI models offer for this scenario?

- A. Predictability of outputs
- B. Adaptability
- C. Less sensitivity to changes in inputs
- D. Explainability

Answer: B

Q187. A financial company wants to flag all credit card activity as possibly fraudulent or non-fraudulent based on transaction data. Which type of ML model meets these requirements?

- A. Regression
- B. Diffusion
- C. Binary classification

D. Multi-class classification

Answer: C

Q188. A company is designing a customer service chatbot by using a fine-tuned large language model (LLM). The company wants to ensure that the chatbot uses responsible AI characteristics.

Select the correct responsible AI characteristic from the following list for each application design action. Each responsible AI characteristic should be selected one time or not at all.

| | |
|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| Anonymize personal information during training data preparation | <div>Select... Select... Governance Privacy and security Safety Transparency</div> |
| Design the customer service chatbot to provide explainable decisions | <div>Select... Select... Governance Privacy and security Safety Transparency</div> |
| Use Amazon Bedrock Guardrails to prevent harmful output and misuse of the chatbot | <div>Select... Select... Governance Privacy and security Safety Transparency</div> |

Answer:

| | |
|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| Anonymize personal information during training data preparation | <div>Select... Select... Governance Privacy and security Safety Transparency</div> |
| Design the customer service chatbot to provide explainable decisions | <div>Select... Select... Governance Privacy and security Safety Transparency</div> |
| Use Amazon Bedrock Guardrails to prevent harmful output and misuse of the chatbot | <div>Select... Select... Governance Privacy and security Safety Transparency</div> |

Q189. A hospital wants to use a generative AI solution with speech-to-text functionality to help improve employee skills in dictating clinical notes. Which AWS service meets these requirements?

- A. Amazon Q Developer
- B. Amazon Polly
- C. Amazon Rekognition

D. AWS HealthScribe

Answer: D

Q190. Which type of AI model makes numeric predictions?

- A. Diffusion
- B. Regression
- C. Transformer
- D. Multi-modal

Answer: B

Q191. A company wants to use Amazon SageMaker features for various use cases.

Select the correct SageMaker feature from the following list for each use case. Each SageMaker feature should be selected one time or not at all.

Preparing data through a visual interface without using code

| |
|-------------------------|
| Select... |
| Select... |
| SageMaker Canvas |
| SageMaker Feature Store |
| SageMaker Ground Truth |
| SageMaker JumpStart |
| SageMaker Model Monitor |

Finding and using a prebuilt solution for fraud detection

| |
|-------------------------|
| Select... |
| Select... |
| SageMaker Canvas |
| SageMaker Feature Store |
| SageMaker Ground Truth |
| SageMaker JumpStart |
| SageMaker Model Monitor |

Create labeled datasets with human intervention

| |
|-------------------------|
| Select... |
| Select... |
| SageMaker Canvas |
| SageMaker Feature Store |
| SageMaker Ground Truth |
| SageMaker JumpStart |
| SageMaker Model Monitor |

Answer:

Preparing data through a visual interface without using code

| |
|-------------------------|
| Select... |
| Select... |
| SageMaker Canvas |
| SageMaker Feature Store |
| SageMaker Ground Truth |
| SageMaker JumpStart |
| SageMaker Model Monitor |

Finding and using a prebuilt solution for fraud detection

| |
|-------------------------|
| Select... |
| Select... |
| SageMaker Canvas |
| SageMaker Feature Store |
| SageMaker Ground Truth |
| SageMaker JumpStart |
| SageMaker Model Monitor |

Create labeled datasets with human intervention

| |
|-------------------------|
| Select... |
| Select... |
| SageMaker Canvas |
| SageMaker Feature Store |
| SageMaker Ground Truth |
| SageMaker JumpStart |
| SageMaker Model Monitor |

Q192. What is the purpose of vector embeddings in a large language model (LLM)?

- A. Splitting text into manageable pieces of data
- B. Grouping a set of characters to be treated as a single unit
- C. Providing the ability to mathematically compare texts
- D. Providing the count of every word in the input

Answer: C

Q193. A company wants to fine-tune a foundation model (FM) by using AWS services. The company needs to ensure that its data stays private, safe, and secure in the source AWS Region where the data is stored. Which combination of steps will meet these requirements MOST cost-effectively? (Choose two.)

- A. Host the model on premises by using AWS Outposts.
- B. Use the Amazon Bedrock API.
- C. Use AWS PrivateLink and a VPC.
- D. Host the Amazon Bedrock API on premises.
- E. Use Amazon CloudWatch logs and metrics.

Answer: BC

Q194. A financial company uses AWS to host its generative AI models. The company must generate reports to show adherence to international regulations for handling sensitive customer data. Which AWS service meets these requirements?

- A. Amazon Macie

- B. AWS Artifact
- C. AWS Secrets Manager
- D. AWS Config

Answer: B

Q195. A medical company wants to modernize its onsite information processing application. The company wants to use generative AI to respond to medical questions from patients. Which AWS service should the company use to ensure responsible AI for the application?

- A. Guardrails for Amazon Bedrock
- B. Amazon Inspector
- C. Amazon Rekognition
- D. AWS Trusted Advisor

Answer: A

Q196. Which metric is used to evaluate the performance of foundation models (FMs) for text summarization tasks?

- A. F1 score
- B. Bilingual Evaluation Understudy (BLEU) score
- C. Accuracy
- D. Mean squared error (MSE)

Answer: B

Q197. What is the benefit of fine-tuning a foundation model (FM)?

- A. Fine-tuning reduces the FM's size and complexity and enables slower inference.
- B. Fine-tuning uses specific training data to retrain the FM from scratch to adapt to a specific use case.
- C. Fine-tuning keeps the FM's knowledge up to date by pre-training the FM on more recent data.
- D. Fine-tuning improves the performance of the FM on a specific task by further training the FM on new labeled data.

Answer: D

Q198. A company wants to improve its chatbot's responses to match the company's desired tone. The company has 100 examples of high-quality conversations between customer service agents and customers. The company wants to use this data to incorporate company tone into the chatbot's responses. Which solution meets these requirements?

- A. Use Amazon Personalize to generate responses.
- B. Create an Amazon SageMaker HyperPod pre-training job.
- C. Host the model by using Amazon SageMaker. Use TensorRT for large language model (LLM) deployment.
- D. Create an Amazon Bedrock fine-tuning job.

Answer: D

Q199. An ecommerce company is using a chatbot to automate the customer order submission process. The chatbot is powered by AI and is available to customers directly from the company's website 24 hours a day, 7 days a week. Which option is an AI system input vulnerability that the company needs to resolve before the chatbot is made available?

- A. Data leakage
- B. Prompt injection
- C. Large language model (LLM) hallucinations
- D. Concept drift

Answer: B

Q200. A social media company wants to prevent users from posting discriminatory content on the company's application. The company wants to use Amazon Bedrock as part of the solution.

How can the company use Amazon Bedrock to meet these requirements?

- A. Give users the ability to interact based on user preferences.
- B. Block interactions related to predefined topics.
- C. Restrict user conversations to predefined topics.
- D. Provide a variety of responses to select from for user engagement.

Answer: B