Q1. 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

Q2. A law firm wants to build an AI application by using large language models (LLMs). The application will read legal documents and extract key points from the documents. Which solution meets these requirements?

A. Build an automatic named entity recognition system.
B. Create a recommendation engine.
C. Develop a summarization chatbot.
D. Develop a multi-language translation system.

Answer: C

Q3. 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

Q4. A company has built an image classification model to predict plant diseases from photos of plant leaves. The company wants to evaluate how many images the model classified correctly. Which evaluation metric should the company use to measure the model's performance?

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

Answer: B

Q5. A company is using a pre-trained large language model (LLM) to build a chatbot for product recommendations. The company needs the LLM outputs to be short and written in a specific language. Which solution will align the LLM response quality with the company's expectations?

A. Adjust the prompt.
B. Choose an LLM of a different size.
C. Increase the temperature.
D. Increase the Top K value.

Answer: A

Q6. 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: C

Q7. 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

Q8. A company is building a solution to generate images for protective eyewear. The solution must have high accuracy and must minimize the risk of incorrect annotations. Which solution will meet these requirements?

A. Human-in-the-loop validation by using Amazon SageMaker Ground Truth Plus
B. Data augmentation by using an Amazon Bedrock knowledge base
C. Image recognition by using Amazon Rekognition
D. Data summarization by using Amazon QuickSight

Answer: A

Q9. 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 with the 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

Q10. 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

Q11. 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

Q12. 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

Q13. A financial institution is using Amazon Bedrock to develop an AI application. The application is hosted in a VPC. To meet regulatory compliance standards, the VPC is not allowed access to any internet traffic. Which AWS service or feature will meet these requirements?

A. AWS PrivateLink
B. Amazon Macie
C. Amazon CloudFront
D. Internet gateway

Answer: A

Q14. 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

Q15. Which metric measures the runtime efficiency of operating AI models?

A. Customer satisfaction score (CSAT)
B. Training time for each epoch
C. Average response time
D. Number of training instances

Answer: C

Q16. A company is building a contact center application and wants to gain insights from customer conversations. The company wants to analyze and extract key information from the audio of the customer calls. Which solution meets these requirements?

A. Build a conversational chatbot by using Amazon Lex.
B. Transcribe call recordings by using Amazon Transcribe.
C. Extract information from call recordings by using Amazon SageMaker Model Monitor.
D. Create classification labels by using Amazon Comprehend.

Answer: B

Q17. 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

Q18. 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

Q19. A company uses a foundation model (FM) from Amazon Bedrock for an AI search tool. The company wants to fine-tune the model to be more accurate by using the company's data. Which strategy will successfully fine-tune the model?

A. Provide labeled data with the prompt field and the completion fielD.
B. Prepare the training dataset by creating a .txt file that contains multiple lines in .csv format.
C. Purchase Provisioned Throughput for Amazon Bedrock.
D. Train the model on journals and textbooks.

Answer: A

Q20. A company wants to use AI to protect its application from threats. The AI solution needs to check if an IP address is from a suspicious source. Which solution meets these requirements?

A. Build a speech recognition system.
B. Create a natural language processing (NLP) named entity recognition system.
C. Develop an anomaly detection system.
D. Create a fraud forecasting system.

Answer: C

Q21. 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

Q22. Which option is a use case for generative AI models?

A. Improving network security by using intrusion detection systems
B. Creating photorealistic images from text descriptions for digital marketing
C. Enhancing database performance by using optimized indexing
D. Analyzing financial data to forecast stock market trends

Answer: B

Q23. A company wants to build a generative AI application by using Amazon Bedrock and needs to choose a foundation model (FM). 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

Q24. A company wants to make a chatbot to help customers. The chatbot will help solve technical problems without human intervention. The company chose a foundation model (FM) for the chatbot. The chatbot needs to produce responses that adhere to company tone.
Which solution meets these requirements?

A. Set a low limit on the number of tokens the FM can produce.
B. Use batch inferencing to process detailed responses.
C. Experiment and refine the prompt until the FM produces the desired responses.
D. Define a higher number for the temperature parameter.

Answer: C

Q25. 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.

Answer: A

Q26. A security company is using Amazon Bedrock to run foundation models (FMs). The company wants to ensure that only authorized users invoke the models. The company needs to identify any unauthorized access attempts to set appropriate AWS Identity and Access Management (IAM) policies and roles for future iterations of the FMs. Which AWS service should the company use to identify unauthorized users that are trying to access Amazon Bedrock?

A. AWS Audit Manager
B. AWS CloudTrail
C. Amazon Fraud Detector
D. AWS Trusted Advisor

Answer: B

Q27. A company has developed an ML model for image classification. The company wants to deploy the model to production so that a web application can use the model. The company needs to implement a solution to host the model and serve predictions without managing any of the underlying infrastructure. Which solution will meet these requirements?

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

Answer: A

Q28. 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: B

Q29. 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

Q30. A company is using the Generative AI Security Scoping Matrix to assess security responsibilities for its solutions. The company has identified four different solution scopes based on the matrix. Which solution scope gives the company the MOST ownership of security responsibilities?

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

Answer: D

Q31. 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

Q32. A company wants to create an application by using Amazon Bedrock. The company has a limited budget and prefers flexibility without long-term commitment. Which Amazon Bedrock pricing model meets these requirements?

A. On-Demand
B. Model customization
C. Provisioned Throughput
D. Spot Instance

Answer: A

Q33. Which AWS service or feature can help an AI development team quickly deploy and consume a foundation model (FM) within the team's VPC?

A. Amazon Personalize
B. Amazon SageMaker JumpStart
C. PartyRock, an Amazon Bedrock Playground
D. Amazon SageMaker endpoints

Answer: B

Q34. How can companies use large language models (LLMs) securely on Amazon Bedrock?

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

Answer: A

Q35. 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

Q36. A company built a deep learning model for object detection and deployed the model to production. Which AI process occurs when the model analyzes a new image to identify objects?

A. Training
B. Inference
C. Model deployment
D. Bias correction

Answer: B

Q37. 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

Q38. A company is implementing the Amazon Titan foundation model (FM) by using Amazon Bedrock. The company needs to supplement the model by using relevant data from the company's private data sources. Which solution will meet this requirement?

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

Answer: C

Q39. A medical company is customizing a foundation model (FM) for diagnostic purposes. The company needs the model to be transparent and explainable to meet regulatory requirements.
Which solution will meet these requirements?

A. Configure the security and compliance by using Amazon Inspector.
B. Generate simple metrics, reports, and examples by using Amazon SageMaker Clarify.
C. Encrypt and secure training data by using Amazon Macie.
D. Gather more datA. Use Amazon Rekognition to add custom labels to the data.

Answer: B

Q40. A company wants to deploy a conversational chatbot to answer customer questions. The chatbot is based on a fine-tuned Amazon SageMaker JumpStart model. The application must comply with multiple regulatory frameworks. Which capabilities can the company show compliance for? (Choose two.)

A. Auto scaling inference endpoints
B. Threat detection
C. Data protection
D. Cost optimization
E. Loosely coupled microservices

Answer: BC

Q41. A company is training a foundation model (FM). The company wants to increase the accuracy of the model up to a specific acceptance level. Which solution will meet these requirements?

A. Decrease the batch size.
B. Increase the epochs.
C. Decrease the epochs.
D. Increase the temperature parameter.

Answer: B

Q42. 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

Q43. Which functionality does Amazon SageMaker Clarify provide?

A. Integrates a Retrieval Augmented Generation (RAG) workflow
B. Monitors the quality of ML models in production
C. Documents critical details about ML models
D. Identifies potential bias during data preparation

Answer: D

Q44. 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

Q45. An ecommerce company wants to build a solution to determine customer sentiments based on written customer reviews of products. Which AWS services meet these requirements? (Choose two.)

A. Amazon Lex
B. Amazon Comprehend
C. Amazon Polly
D. Amazon Bedrock
E. Amazon Rekognition

Answer: B, D

Q46. A company wants to use large language models (LLMs) with Amazon Bedrock to develop a chat interface for the company's product manuals. The manuals are stored as PDF files. Which solution meets these requirements MOST cost-effectively?

A. Use prompt engineering to add one PDF file as context to the user prompt when the prompt is submitted to Amazon Bedrock.
B. Use prompt engineering to add all the PDF files as context to the user prompt when the prompt is submitted to Amazon Bedrock.
C. Use all the PDF documents to fine-tune a model with Amazon Bedrock. Use the fine-tuned model to process user prompts.
D. Upload PDF documents to an Amazon Bedrock knowledge base. Use the knowledge base to provide context when users submit prompts to Amazon Bedrock.

Answer: D

Q47. A social media company wants to use a large language model (LLM) for content moderation. The company wants to evaluate the LLM outputs for bias and potential discrimination against specific groups or individuals. Which data source should the company use to evaluate the LLM outputs with the LEAST administrative effort?

A. User-generated content
B. Moderation logs
C. Content moderation guidelines
D. Benchmark datasets

Answer: D

Q48. A company wants to use a pre-trained generative AI model to generate content for its marketing campaigns. The company needs to ensure that the generated content aligns with the company's brand voice and messaging requirements. Which solution meets these requirements?

A. Optimize the model's architecture and hyperparameters to improve the model's overall performance.
B. Increase the model's complexity by adding more layers to the model's architecture.
C. Create effective prompts that provide clear instructions and context to guide the model's generation.
D. Select a large, diverse dataset to pre-train a new generative model.

Answer: C

Q49. 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? (Choose 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

Q50. A company is using an Amazon Bedrock base model to summarize documents for an internal use case. The company trained a custom model to improve the summarization quality.
Which action must the company take to use the custom model through Amazon Bedrock?

A. Purchase Provisioned Throughput for the custom model.
B. Deploy the custom model in an Amazon SageMaker endpoint for real-time inference.
C. Register the model with the Amazon SageMaker Model Registry.
D. Grant access to the custom model in Amazon Bedrock.

Answer: D