



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

119. Which of the following is NOT a typical business metric for evaluating generative AI applications?	Model parameter count
5. A company is using a pre-trained large language model (LLM) to build a chatbot for product recommendations. The company needs the LLM's responses to be brief and in a specific language. Which solution best aligns the response quality with the company's expectations?	Adjust the prompt
54. A digital devices company wants to develop an interactive application for children that generates new stories based on classic tales. The application must ensure that the content is appropriate for children. Which AWS service or feature best meets this requirement?	C. Guardrails for Amazon Bedrock
50. A company is using a base model from Amazon Bedrock to summarize documents for an internal use case. It trained a custom model to improve summary quality. How should the company deploy the custom model via Amazon Bedrock?	Deploy the custom model on an SageMaker Endpoint for real-time inference.
92. What does AUC stand for in model performance metrics?	Area Under the Curve
3. A company wants to classify human genes into 20 categories based on their characteristics. The company needs an ML algorithm that can document how the model's internal decision-making affects its output. Which ML algorithm meets these requirements?	Decision Trees
46. A company wants to develop a chat interface for its product manuals stored as PDFs using a large language model (LLM) on Amazon Bedrock. Which solution is the most cost-effective?	! Using a knowledge base to supply context is the most cost-effective method since it avoids the need for extensive fine-tuning or processing each prompt with all PDFs.
What is the primary purpose of a data catalog in AI governance?	INCORRECT. The correct answer is: b) To organize and inventory data assets.. A data catalog provides a structured, searchable repository of data assets, making it easier to manage, discover, and govern data used in AI systems
134. What is a chain-of-thought prompt?	CORRECT! A chain-of-thought prompt is designed to make the model articulate intermediate reasoning steps before delivering the final answer, which can improve performance on complex tasks.
131. What is Retrieval-Augmented Generation (RAG)?	CORRECT! RAG combines information retrieval with text generation, allowing the model to incorporate external knowledge into its outputs, thereby improving relevance and accuracy.
108. What is the key advantage of generative AI's adaptability?	CORRECT! Generative AI is highly adaptable and can be fine-tuned for many different tasks across various domains without needing to be completely retrained from scratch.
127. Which of the following is NOT a typical use case for generative AI in customer service?	CORRECT! Generative AI is mainly used for tasks such as text generation, chatbots, and automated responses. Physical robot assistants typically involve robotics and are not a common use case for generative AI.
144. What is the primary objective of responsible model selection?	CORRECT! Responsible model selection involves weighing technical performance alongside ethical, social, and environmental concerns.
150. What is the primary objective of negative prompts in prompt engineering?	INCORRECT. The correct answer is: b) To instruct the model on what to avoid in its output.. Negative prompts specify what the model should not generate, helping prevent undesired or harmful outputs.
132. Which AWS service is well-suited for storing embeddings in a vector database?	CORRECT! Amazon OpenSearch Service supports k-NN (k-nearest neighbors) search, making it ideal for storing and querying vector embeddings used in generative AI applications.
139. What is the primary objective of instruction fine-tuning in foundation models?	INCORRECT. The correct answer is: a) To teach the model to follow specific instructions.. Instruction fine-tuning involves training the model with input-output pairs that are formatted as instructions, helping it learn to respond appropriately to user commands.
58. A company is developing a customer support chatbot that improves its responses by learning from past interactions and online resources. Which AI learning strategy enables this self-improvement?	CORRECT! Reinforcement learning enables the chatbot to adapt its behavior based on feedback, thereby continuously improving its responses over time. HOMENext



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

109. What is a hallucination in the context of generative AI?	INCORRECT. The correct answer is: c) An output that is fabricated or factually incorrect despite appearing plausible. Hallucination refers to when a model produces responses that sound plausible but are incorrect or made-up
112. Which of the following is NOT a stage in the lifecycle of a foundation model?	CORRECT! The typical lifecycle of a foundation model includes data selection, pre-training, fine-tuning, evaluation, and deployment—marketing is not part of the technical lifecycle.
145. Which of the following is NOT typically a consideration for datasets in responsible AI?	CORRECT! While size is important, responsible AI emphasizes data quality, provenance, and balanced representation over simply having a very large dataset.
Which of the following is NOT a typical regulatory compliance standard for AI systems?	INCORRECT. The correct answer is: c) HTML. HTML is a markup language for web pages, not a regulatory standard. ISO, SOC, and algorithm accountability laws are examples of compliance standards relevant to AI.
201. Which of the following is NOT typically a component of a data governance strategy for AI systems?	INCORRECT. The correct answer is: c) Public data sharing policies.. A data governance strategy is focused on managing and securing data, not on making it publicly available without restrictions. HOMEPreviousNext
206. Which AWS services or resources should a company use to continuously monitor the performance of a fraud detection model in production and incorporate human review when necessary? (Select TWO.)	INCORRECT. The correct answer is: a) Amazon SageMaker Model Monitor, b) Amazon A2I (Amazon Augmented AI). Amazon SageMaker Model Monitor tracks model performance and data drift continuously, while Amazon A2I integrates human review for low-confidence predictions or periodic audits.
84. A company uses SageMaker Studio notebooks for developing and training ML models, with data stored in an Amazon S3 bucket. How should the company manage data flow between S3 and SageMaker Studio?	The correct answer is: C. Configure SageMaker to use a VPC with an S3 endpoint.. Setting up a VPC with an S3 endpoint enables secure and efficient data transfer between SageMaker Studio and S3.
202. What is the primary objective of AWS CloudTrail in AI governance?	CORRECT! AWS CloudTrail logs all API calls and account activities, providing a critical audit trail for compliance and security purposes.
171. Which of the following is NOT a typical method for improving model interpretability?	INCORRECT. The correct answer is: c) Increasing the model's size.. Larger models are generally harder to interpret, whereas simpler models, feature importance metrics, and clear explanations enhance interpretability.
87. Which of the following is NOT a type of machine learning?	INCORRECT. The correct answer is: d) Diagnostic learning. The main types of machine learning are supervised, unsupervised, and reinforcement learning; 'diagnostic learning' is not a recognized category. HOMEPreviousNext
104. What is the primary goal of model monitoring in production?	INCORRECT. The correct answer is: c) To detect issues such as model drift or data drift. Monitoring ensures that models continue to perform as expected over time by detecting changes in data or performance that may require intervention.
69. An education provider is developing a Q&A application using a generative AI model to explain complex concepts. The model should automatically adjust its response style based on the age range provided for the user. Which solution requires the least implementation effort?	B. Include a role description in the prompt context specifying the target age range for the response.. Adding a role description to the prompt is a lightweight and effective method to guide the model's response style according to the user's age without additional training.
162. Which AWS service is designed to detect and monitor biases in ML models?	CORRECT! Amazon SageMaker Clarify offers tools to identify and mitigate biases in both data and models, enhancing transparency and fairness
2. A law firm wants to develop an AI application using large language models (LLMs) that can read legal documents and extract their key points. Which solution meets these requirements?	CORRECT! A summarization chatbot based on an LLM can be trained to read and condense legal documents, extracting only the most critical information.
55. A company wants to generate synthetic data based on its existing data. Which type of model is best suited for this task?	INCORRECT. The correct answer is: A. Generative Adversarial Network (GAN). GANs are specifically designed to generate realistic synthetic data by learning the distribution of the input data.
12. A company wants to leverage generative AI to boost developer productivity and speed up software development by using Amazon Q Developer. What does this service do?	



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

	A. Create software snippets, track references, and manage open-source licenses.. Amazon Q Developer is specifically designed to help create code snippets, manage open-source licenses, and track references, streamlining software development.
62. A company built an ML model for fraud detection on its e-commerce platform that has been in production for several months. To ensure it continues to perform accurately and reliably, which AWS services or resources should be used to monitor its performance and incorporate human review when necessary? (Select TWO.)	NCORRECT. The correct answer is: a) Amazon SageMaker Model Monitor, b) Amazon A2I (Amazon Augmented AI). Amazon SageMaker Model Monitor tracks deployed models for issues like data drift, while Amazon A2I allows for the easy integration of human review for low-confidence predictions or periodic audits.
185. What is data lineage in the context of AI governance?	CORRECT! Data lineage documents the history of data from its source through all transformations, which is vital for governance and auditing purposes.
1) A machine learning team has several large CSV datasets in Amazon S3. Historically, models built with the Amazon SageMaker Linear Learner algorithm have taken hours to train on similar-sized datasets. The team's leaders need to accelerate the training process. What can a machine learning specialist do to address this concern?	1) A - Amazon SageMaker Pipe mode streams the data directly to the container, which improves the performance of training jobs. (Refer to this link for supporting information.) In Pipe mode, your training job streams data directly from Amazon S3. Streaming can provide faster start times for training jobs and better throughput. With Pipe mode, you also reduce the size of the Amazon EBS volumes for your training instances. B would not apply in this scenario. C is a streaming ingestion solution, but is not applicable in this scenario. D transforms the data structure.
2) A term frequency-inverse document frequency (tf-idf) matrix using both unigrams and bigrams is built from a text corpus consisting of the following two sentences: 1. Please call the number below. 2. Please do not call us.	2) A - There are 2 sentences, 8 unique unigrams, and 8 unique bigrams, so the result would be (2,16). The phrases are "Please call the number below" and "Please do not call us." Each word individually (unigram) is "Please," "call," "the," "number," "below," "do," "not," and "us." The unique bigrams are "Please call," "call the," "the number," "number below," "Please do," "do not," "not call," and "call us." The tf-idf vectorizer is described at this link.
3) A company is setting up a system to manage all of the datasets it stores in Amazon S3. The company would like to automate running transformation jobs on the data and maintaining a catalog of the metadata concerning the datasets. The solution should require the least amount of setup and maintenance. Which solution will allow the company to achieve its goals? A) Create an Amazon EMR cluster with Apache Hive installed	3) B - AWS Glue is the correct answer because this option requires the least amount of setup and maintenance since it is serverless, and it does not require management of the infrastructure. Refer to this link for supporting information. A, C, and D are all solutions that can solve the problem, but require more steps for configuration, and require higher operational overhead to run and maintain.
4) A data scientist is working on optimizing a model during the training process by varying multiple parameters. The data scientist observes that, during multiple runs with identical parameters, the loss function converges to different, yet stable, values. What should the data scientist do to improve the training process?	4) B - It is most likely that the loss function is very curvy and has multiple local minima where the training is getting stuck. Decreasing the batch size would help the data scientist stochastically get out of the local minima saddles. Decreasing the learning rate would prevent overshooting the global loss function minimum. Refer to the paper at this link for an explanation.
5) A data scientist is evaluating different binary classification models. A false positive result is 5 times more expensive (from a business perspective) than a false negative result. The models should be evaluated based on the following criteria: 1) Must have a recall rate of at least 80% 2) Must have a false positive rate of 10% or less 3) Must minimize business costs After creating each binary classification model, the data scientist generates the corresponding confusion-matrix. Which confusion matrix represents the model that satisfies the requirements?	5) D - The following calculations are required: TP = True Positive FP = False Positive FN = False Negative TN = True Negative FN = False Negative Recall = TP / (TP + FN) False Positive Rate (FPR) = FP / (FP + TN) Cost = 5 * FP + FN Options C and D have a recall greater than 80% and an FPR less than 10%, but D is the most cost effective. For supporting information, refer to this link.
6) A data scientist uses logistic regression to build a fraud detection model. While the model accuracy is 99%, 90% of the fraud cases are not detected by the model. What action will definitely help the model detect more than 10% of fraud cases?	6) B - Decreasing the class probability threshold makes the model more sensitive and, therefore, marks more cases as the positive class, which is fraud in this case. This will increase the likelihood of fraud detection. However, it comes at the price of lowering precision. This is covered in the Discussion section of the paper at this link.
7) A company is interested in building a fraud detection model. Currently, the data scientist does not have a sufficient amount of information due to the low number of fraud cases. Which method is MOST likely to detect the GREATEST number of valid fraud cases? A) Oversampling using bootstrapping B) Undersampling C) Oversampling using SMOTE D) Class weight adjustment	7) C - With datasets that are not fully populated, the Synthetic Minority Over-sampling Technique (SMOTE) adds new information by adding synthetic data points to the minority class. This technique would be the most effective in this scenario. Refer to Section 4.2 at this link for supporting information.
8) A machine learning engineer is preparing a data frame for a supervised learning task with the Amazon SageMaker Linear Learner algorithm. The ML engineer notices the target	8) D - Use supervised learning to predict missing values based on the values of other features. Different supervised learning approaches might have different performances, but



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

label classes are highly imbalanced and multiple feature columns contain missing values. The proportion of missing values across the entire data frame is less than 5%. What should the ML engineer do to minimize bias due to missing values?	any properly implemented supervised learning approach should provide the same or better approximation than mean or median approximation, as proposed in responses A and C. Supervised learning applied to the imputation of missing values is an active field of research. Refer to this link for an example.
9) A company has collected customer comments on its products, rating them as safe or unsafe using decision trees. The training dataset has the following features: id, date, full review, full review summary, and a binary safe/unsafe tag. During training, any data sample with missing features was dropped. In a few instances, the test set was found to be missing the full review text field. For this use case, which is the most effective course of action to address test data samples with missing features?	9) B - In this case, a full review summary usually contains the most descriptive phrases of the entire review and is a valid stand-in for the missing full review text field. For supporting information, refer to page 1627 at this link, and this link and this link
10) An insurance company needs to automate claim compliance reviews because human reviews are expensive and error-prone. The company has a large set of claims and a compliance label for each. Each claim consists of a few sentences in English, many of which contain complex related information. Management would like to use Amazon SageMaker built-in algorithms to design a machine learning supervised model that can be trained to read each claim and predict if the claim is compliant or not. Which approach should be used to extract features from the claims to be used as inputs for the downstream supervised task?	D) Apply Amazon SageMaker Object2Vec to claims in the training set. Send the derived features space as inputs for the downstream supervised task.
Case Study -A company is building a web-based AI application by using Amazon SageMaker. The application will provide the following capabilities and features: ML experimentation, training, a central model registry, model deployment, and model monitoring.The application must ensure secure and isolated use of training data during the ML lifecycle. The training data is stored in Amazon S3.The company needs to use the central model registry to manage different versions of models in the application.Which action will meet this requirement with the LEAST operational overhead? A. Create a separate Amazon Elastic Container Registry (Amazon ECR) repository for each model. B. Use Amazon Elastic Container Registry (Amazon ECR) and unique tags for each model version. C. Use the SageMaker Model Registry and model groups to catalog the models. Most Voted D. Use the SageMaker Model Registry and uniqu	C.
Case Study -A company is building a web-based AI application by using Amazon SageMaker. The application will provide the following capabilities and features: ML experimentation, training, a central model registry, model deployment, and model monitoring.The application must ensure secure and isolated use of training data during the ML lifecycle. The training data is stored in Amazon S3.The company is experimenting with consecutive training jobs.How can the company MINIMIZE infrastructure startup times for these jobs? A. Use Managed Spot Training. B. Use SageMaker managed warm pools. Most Voted C. Use SageMaker Training Compiler. D. Use the SageMaker distributed data parallelism (SMDDP) library.	Use SageMaker managed warm pools
A company is building a web-based AI application by using Amazon SageMaker. The application will provide the following capabilities and features: ML experimentation, training, a central model registry, model deployment, and model monitoring.The application must ensure secure and isolated use of training data during the ML lifecycle. The training data is stored in Amazon S3.The company must implement a manual approval-based workflow to ensure that only approved models can be deployed[...] A. Use SageMaker Experiments to facilitate the approval process during model registration. B. Use SageMaker ML Lineage Tracking on the central model registry. Create tracking entities for the approval process.	D. Use SageMaker Pipelines. When a model version is registered, use the AWS SDK to change the approval status to "Approved." Most Voted



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

C. Use SageMaker Model Monitor to evaluate the performance of the model and to manage the approval. D. Use SageMaker Pipelines. When a model version is registered, use the AWS SDK to change the approval status to "Approved."	
Case Study -A company is building a web-based AI application by using Amazon SageMaker. The application will provide the following capabilities and features: ML experimentation, training, a central model registry, model deployment, and model monitoring.The application must ensure secure and isolated use of training data during the ML lifecycle. The training data is stored in Amazon S3.The company needs to run an on-demand workflow to monitor bias drift for models that are deployed to real-time endpoints from the application.Which action will meet this requirement? A. Configure the application to invoke an AWS Lambda function that runs a SageMaker Clarify job. B. Invoke an AWS Lambda function to pull the sagemaker-model-monitor-analyzer built-in SageMaker image. C. Use AWS Glue Data Quality to monitor bias. D. Use SageMaker notebooks to compare the bias.	A. Configure the application to invoke an AWS Lambda function that runs a SageMaker Clarify job. Most Voted
156. Which of the following is NOT a typical characteristic of responsible AI?	INCORRECT. The correct answer is: c) Profitability. Responsible AI focuses on ethical principles such as fairness, transparency, and inclusivity rather than on profitability alone. HOMENext
172. What is the primary purpose of data governance in AI?	CORRECT! Data governance involves creating a structured, searchable catalog of data assets, which facilitates better management and understanding of the data used in AI systems.
49. A lending company is developing a generative AI solution to offer discounts to new applicants based on specific business criteria. To use the AI model responsibly and minimize bias, which actions should the company take? (Choose two.)	A. Detect imbalances or disparities in the data., C. Evaluate the model's behavior to ensure transparency for stakeholders.. Detecting data imbalances helps prevent bias during training, while evaluating the model's behavior provides transparency and accountability, both of which are essential for responsible AI use.
51. A company needs to select an Amazon Bedrock model for internal use that generates responses in a style preferred by its employees. What should the company do?	INCORRECT. The correct answer is: B. Evaluate models using human evaluators and custom prompt datasets.. Using human evaluation with custom prompts ensures the chosen model generates responses that align with the specific style and preferences of the employees.
126. What is the primary purpose of embeddings in generative AI?	CORRECT! Embeddings transform data into vector representations in high-dimensional space, capturing semantic relationships and enabling models to process and understand complex data.
158. What is a key consideration regarding the environmental impact when selecting a model?	CORRECT! Responsible model selection involves evaluating not just performance, but also the environmental footprint associated with the model's development and deployment.
52. A university student is copying generative AI content to write essays. Which responsible AI challenge does this scenario represent?	INCORRECT. The correct answer is: C. Plagiarism. Plagiarism occurs when AI-generated content is used without proper attribution or originality, posing ethical challenges in responsible AI use.
70. Which strategy is used to evaluate the accuracy of a foundation model (FM) for image classification?	The correct answer is: B. Measure the model's accuracy against a predefined benchmark dataset.. Evaluating accuracy using a benchmark dataset is a standard method to assess the performance of an image classification model.
91. Which of the following is NOT a typical stage in an ML pipeline?	INCORRECT. The correct answer is: d) Customer acquisition. An ML pipeline generally covers data collection, preprocessing, modeling, and evaluation. Customer acquisition is a business process, not a technical ML stage.
44. A company developed a new model to predict the prices of certain items. While the model performed well on the training data, its performance dropped significantly in production. What should the company do to mitigate this issue?	CORRECT! Increasing the training data volume can help the model generalize better and reduce overfitting, leading to improved production performance.



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

83. A company wants to evaluate the cost of using an LLM for inference on Amazon Bedrock. Which factor directly influences the cost?	INCORRECT. The correct answer is: A. The number of tokens processed.. Inference cost is directly proportional to the number of tokens processed, as each token consumes computational resources.
160. Which of the following is NOT typically a business metric for evaluating AI applications?	INCORRECT. The correct answer is: d) Model parameter count. Model parameter count is a technical metric reflecting model complexity, while business metrics focus on financial and customer impact.
113. What is the main advantage of using AWS generative AI services for building applications?	CT. The correct answer is: b) They provide a lower barrier to entry.. AWS generative AI services, such as Amazon Bedrock, offer pre-trained models and user-friendly tools that simplify development, reducing the barrier to entry for businesses.
90. What is the primary objective of exploratory data analysis (EDA) in the ML lifecycle?	INCORRECT. The correct answer is: c) To understand the characteristics of the data. EDA involves analyzing and visualizing data to understand its main features, detect patterns, and identify anomalies before model training.
106. In the context of generative AI, what is a token?	A unit of text processed by the model
148. Which AWS service is best suited for storing embeddings in a relational database?	c) Amazon Aurora. Amazon Aurora is a relational database (compatible with MySQL and PostgreSQL) that supports storing and querying vector embeddings, making it suitable for such tasks.
181. Which AWS service is primarily used for managing access and permissions for AI systems?	IAM
141. Which of the following is NOT a typical benefit of using open-source models for transparency?	c) A guarantee of perfect performance.. While open-source models offer transparency, code inspection, and community contributions, they do not guarantee flawless performance.
138. What is a potential risk of using zero-shot learning in prompt engineering?	a) The model may perform poorly on tasks it wasn't explicitly trained for.. Zero-shot learning uses a model on a task without explicit training for it; if the task is too different, performance may suffer.
149. What is a key consideration when evaluating whether a foundation model meets business objectives?	CORRECT! When evaluating a model for business use, metrics such as user engagement, customer satisfaction, and ROI are critical alongside technical performance.
146. What is prompt modeling in the context of prompt engineering?	CORRECT! Prompt modeling involves designing templates that can be reused with different inputs, streamlining the process of prompt engineering.
107. Which of the following is NOT a typical use case for generative AI models?	c) Data encryption. Generative AI is used to create new content such as images, text, or code. Data encryption is not a typical use case for generative models.
140. What is BERTScore used for in evaluating foundation models?	CORRECT! BERTScore uses contextual embeddings from BERT to assess the similarity between generated text and a reference, thereby measuring the quality of the output.
157. What is the primary purpose of Guardrails for Amazon Bedrock?	CORRECT! Guardrails for Amazon Bedrock help implement safeguards that ensure the AI system operates responsibly, adhering to ethical and security standards.
136. What is the main purpose of ROUGE in evaluating foundation models?	b) To evaluate the quality of generated summaries. ROUGE is a set of metrics used to assess the quality of summaries generated by comparing them with human-written reference summaries.
133. What is the primary purpose of adjusting the temperature parameter during inference?	CORRECT! The temperature parameter modulates the randomness of the output. A higher temperature results in more varied responses, while a lower temperature yields more deterministic output
122. What is a key benefit of using generative AI models in terms of development speed?	: b) They offer a faster time-to-market.. AWS generative AI services, such as Amazon Bedrock, speed up development by providing pre-trained models and easy-to-use tools, reducing overall development time.
154. What is a key benefit of using transfer learning for personalizing a foundation model?	CORRECT! Transfer learning enables a model trained on one task to be adapted for a related task, thus improving performance with less additional training data.



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

155. What is the primary purpose of a model's latent space in prompt engineering?	CORRECT! The latent space is a mathematical representation of the model's internal features and knowledge, which can be manipulated to influence output in a controlled way
147. What is the main advantage of using few-shot learning in prompt engineering?	CORRECT! Few-shot learning lets the model generalize from just a few examples provided in the prompt, which is particularly useful when limited training data is available.
57. A retailer installed a security camera and uses an ML model to analyze the images for potential theft. The model disproportionately flags people from a specific ethnic group. Which type of bias is affecting the model's output?	CORRECT! Sampling bias occurs when the training data does not adequately represent all groups, resulting in disproportionate outcomes for certain demographics.
19. A company is using a foundation model (FM) from Amazon Bedrock for an AI search tool and wants to fine-tune it with its own data to improve accuracy. What is the best approach?	A. Provide labeled data that includes both a prompt and a corresponding completion.. Fine-tuning a foundation model requires labeled data with prompts and expected completions to guide the model's adjustment to company-specific data.
42. A company is developing a chatbot based on a large language model (LLM) to answer customer questions and wants to reduce the workload for call center staff. Which business metric should be used to evaluate the chatbot's impact?	B. Average call duration. A decrease in average call duration indicates that the chatbot is effectively reducing the need for human intervention in customer support.
130. Which AWS service helps developers quickly get started with pre-trained generative AI models?	CORRECT! Amazon SageMaker JumpStart provides access to pre-trained models, notebooks, and ready-to-use solutions, making it easy for developers to begin building ML applications.
Question 28 of 65	
32. A company wants to create an application using Amazon Bedrock but has a limited budget and prefers flexibility without long-term commitments. Which Amazon Bedrock pricing model meets these requirements?	A. On-demand. The on-demand pricing model lets you pay only for the resources you use without any long-term commitments, making it ideal for companies with limited budgets.
33. Which AWS service or resource can help an AI development team quickly deploy and utilize a foundation model (FM) within their VPC?	INCORRECT. The correct answer is: B. Amazon SageMaker JumpStart. Amazon SageMaker JumpStart provides pre-trained models and ready-to-use notebooks, enabling rapid deployment and testing within a VPC.
114. What is prompt engineering in the context of generative AI?	CORRECT! Prompt engineering involves creating and refining the instructions given to a model to obtain the desired output, making it a key aspect of working with generative AI.
26. A security company is using Amazon Bedrock to run foundation models (FMs). To ensure that only authorized users can invoke these models, the company needs to detect unauthorized access attempts and set up appropriate IAM policies and roles. Which AWS service should be used?	INCORRECT. The correct answer is: B. AWS CloudTrail. AWS CloudTrail logs all account activity, including unauthorized access attempts, allowing the company to identify and respond to unauthorized users.
29. A company wants to develop a conversational agent using a large language model (LLM) but must prevent manipulation through prompt engineering techniques that could lead to unwanted actions or exposure of sensitive information. What is the best approach?	A. Create a prompt template that teaches the LLM to detect and ignore malicious instructions.. A well-crafted prompt template can help the LLM recognize and avoid attack patterns, significantly reducing the risk of manipulation
128. What is the primary benefit of using AWS generative AI services for application development in terms of speed?	b) They enable a faster time-to-market.. AWS generative AI services accelerate development by providing pre-built models and easy-to-use tools, thereby reducing development time and speeding up time-to-market.
4. A company built an image classification model to predict plant diseases from photos of leaves. The company wants to evaluate how many images the model classifies correctly. Which evaluation metric should be used?	b) Accuracy.. Accuracy measures the proportion of images correctly classified, making it a suitable metric for evaluating the performance of an image classification model.
8. A company is developing a solution to generate images for safety goggles. The solution must be highly accurate and minimize the risk of incorrect annotations. Which solution meets these requirements?	a) Human-in-the-loop validation using Amazon SageMaker Ground Truth Plus.. Amazon SageMaker Ground Truth Plus enables human reviewers to verify and correct annotations, ensuring high accuracy and reducing the likelihood of errors.
118. What is a key consideration when selecting a generative AI model for a business problem?	CORRECT! Choosing a generative AI model should be based on performance metrics such as speed, accuracy, and capacity to handle the specific task, rather than on factors like popularity or origin.



MLA-C01 Practice Quiz (35 + 65 + 14) Questions

Study online at https://quizlet.com/_hyuxwq

37. An AI practitioner built a deep learning model to generate images of people in various professions. He noticed that the input data is biased and that certain attributes influence the generated images, causing bias. Which technique can address this issue?	INCORRECT. The correct answer is: A. Data augmentation for imbalanced classes. Data augmentation for imbalanced classes helps correct biases by ensuring all classes are well-represented during training.
38. A company is implementing the Amazon Titan foundation model (FM) using Amazon Bedrock and needs to supplement it with relevant data from its private sources. Which solution meets this requirement?	C. Create a knowledge base in Amazon Bedrock.. Creating a knowledge base in Amazon Bedrock allows the company to integrate its private data with the foundation model, enhancing its performance without retraining from scratch.
23. A company wants to build a generative AI application using Amazon Bedrock and needs to select a foundation model (FM). The company is interested in knowing how much information can fit in a single prompt. What parameter is this referring to?	Context window
31. An AI practitioner has a database of animal photos and wants to automatically identify and categorize the animals without manual intervention. Which approach should be used?	The correct answer is: A. Object detection. Object detection is ideally suited for automatically identifying and categorizing objects, such as animals in images, without human input. HOMEPrevious
41. A company is training a foundation model (FM) and wants to boost its accuracy to a certain acceptable level. Which approach should be taken?	B. Increase the number of epochs.. Increasing the number of epochs allows the model to learn more thoroughly from the training data, potentially improving accuracy provided that overfitting is controlled.
27. A company developed an ML model for image classification and wants to deploy it so that a web application can make predictions without managing infrastructure. Which solution is best?	CORRECT! Amazon SageMaker Serverless Inference lets you deploy models without managing servers, offering scalable and simplified prediction services.
108. What is the key advantage of generative AI's adaptability?	CORRECT! Generative AI is highly adaptable and can be fine-tuned for many different tasks across various domains without needing to be completely retrained from scratch
110. Which AWS service is specifically designed for developing generative AI applications?	The correct answer is: c) Amazon Bedrock. Amazon Bedrock is a fully managed service that gives you access to foundation models from leading AI startups and Amazon, ideal for generative AI applications.
167. Which of the following is NOT a typical effect of bias in AI systems?	INCORRECT. The correct answer is: b) Improved overall accuracy.. Bias in AI typically leads to unfair outcomes, legal challenges, and loss of trust rather than improved accuracy overall.
171. Which of the following is NOT a typical method for improving model interpretability?	c) Increasing the model's size.. Larger models are generally harder to interpret, whereas simpler models, feature importance metrics, and clear explanations enhance interpretability.
56. An e-commerce company wants to forecast customer demand for memory hardware. Lacking programming expertise and ML knowledge, the company needs a data-driven predictive model that can analyze both internal and external data. Which solution meets these requirements?	Import data into SageMaker Canvas and build ML models by selecting values from the Canvas data.
77. An AI practitioner is using an LLM to generate marketing content. Although the output appears plausible and factual, it is incorrect. What issue is the LLM experiencing?	HALLUCINATION
Question 52 of 65 183. What does the AWS shared responsibility model refer to?	CORRECT! The AWS shared responsibility model delineates what AWS secures (the cloud infrastructure) and what the customer is responsible for (data, applications, etc.).
76. A company is using few-shot prompting with a base model on Amazon Bedrock that currently uses 10 examples in the prompt. The model is invoked once daily and performs well, but the company wants to reduce costs. What is the best solution?	: B. Reduce the number of tokens in the prompt.. Reducing the token count in the prompt can lower processing costs without affecting performance, especially when the model is invoked infrequently.
189. Which AWS service is used for continuous monitoring and evaluation of resources?	AWS Config
205. Which of the following is NOT typically a component of a data lifecycle management strategy?	CORRECT! Data lifecycle management focuses on the controlled creation, storage, retention, and deletion of data rather than on open public sharing.
20. A company wants to use AI to secure its application by verifying whether an IP address originates from a suspicious source. Which solution is most appropriate?	Develop an anomaly detection system