

Oracle

1Z0-1127-25 Exam

Oracle Cloud Infrastructure 2025 Generative AI Professional

**Questions & Answers
Demo**

Version: 6.0

Question: 1

What is the role of temperature in the decoding process of a Large Language Model (LLM)?

- A. To increase the accuracy of the most likely word in the vocabulary
- B. To determine the number of words to generate in a single decoding step
- C. To decide to which part of speech the next word should belong
- D. To adjust the sharpness of probability distribution over vocabulary when selecting the next word

Answer: D

Explanation:

Comprehensive and Detailed In-Depth Explanation=

Temperature is a hyperparameter in the decoding process of LLMs that controls the randomness of word selection by modifying the probability distribution over the vocabulary. A lower temperature (e.g., 0.1) sharpens the distribution, making the model more likely to select the highest-probability words, resulting in more deterministic and focused outputs. A higher temperature (e.g., 2.0) flattens the distribution, increasing the likelihood of selecting less probable words, thus introducing more randomness and creativity. Option D accurately describes this role. Option A is incorrect because temperature doesn't directly increase accuracy but influences output diversity. Option B is unrelated, as temperature doesn't dictate the number of words generated. Option C is also incorrect, as part-of-speech decisions are not directly tied to temperature but to the model's learned patterns.

: General LLM decoding principles, likely covered in OCI 2025 Generative AI documentation under decoding parameters like temperature.

Question: 2

Which statement accurately reflects the differences between these approaches in terms of the number of parameters modified and the type of data used?

- A. Fine-tuning and continuous pretraining both modify all parameters and use labeled, task-specific data.
- B. Parameter Efficient Fine-Tuning and Soft Prompting modify all parameters of the model using

unlabeled data.

C. Fine-tuning modifies all parameters using labeled, task-specific data, whereas Parameter Efficient Fine-Tuning updates a few, new parameters also with labeled, task-specific data.

D. Soft Prompting and continuous pretraining are both methods that require no modification to the original parameters of the model.

Answer: C

Explanation:

Comprehensive and Detailed In-Depth Explanation=

Fine-tuning typically involves updating all parameters of an LLM using labeled, task-specific data to adapt it to a specific task, which is computationally expensive. Parameter Efficient Fine-Tuning (PEFT), such as methods like LoRA (Low-Rank Adaptation), updates only a small subset of parameters (often newly added ones) while still using labeled, task-specific data, making it more efficient. Option C correctly captures this distinction. Option A is wrong because continuous pretraining uses unlabeled data and isn't task-specific. Option B is incorrect as PEFT and Soft Prompting don't modify all parameters, and Soft Prompting typically uses labeled examples indirectly. Option D is inaccurate because continuous pretraining modifies parameters, while Soft Prompting doesn't.

: OCI 2025 Generative AI documentation likely discusses Fine-tuning and PEFT under model customization techniques.

Question: 3

What is prompt engineering in the context of Large Language Models (LLMs)?

A. Iteratively refining the ask to elicit a desired response

B. Adding more layers to the neural network

C. Adjusting the hyperparameters of the model

D. Training the model on a large dataset

Answer: A

Explanation:

Comprehensive and Detailed In-Depth Explanation=

Prompt engineering involves crafting and refining input prompts to guide an LLM to produce desired outputs without altering its internal structure or parameters. It's an iterative process that leverages the model's pre-trained knowledge, making Option A correct. Option B is unrelated, as adding layers pertains to model architecture design, not prompting. Option C refers to hyperparameter tuning (e.g., temperature), not prompt engineering. Option D describes pretraining or fine-tuning, not prompt engineering.

: OCI 2025 Generative AI documentation likely covers prompt engineering in sections on model interaction or inference.

Question: 4

What does the term "hallucination" refer to in the context of Large Language Models (LLMs)?

- A. The model's ability to generate imaginative and creative content
- B. A technique used to enhance the model's performance on specific tasks
- C. The process by which the model visualizes and describes images in detail
- D. The phenomenon where the model generates factually incorrect information or unrelated content as if it were true

Answer: D

Explanation:

Comprehensive and Detailed In-Depth Explanation=

In LLMs, "hallucination" refers to the generation of plausible-sounding but factually incorrect or irrelevant content, often presented with confidence. This occurs due to the model's reliance on patterns in training data rather than factual grounding, making Option D correct. Option A describes a positive trait, not hallucination. Option B is unrelated, as hallucination isn't a performance-enhancing technique. Option C pertains to multimodal models, not the general definition of hallucination in LLMs.

: OCI 2025 Generative AI documentation likely addresses hallucination under model limitations or evaluation metrics.

Question: 5

What does in-context learning in Large Language Models involve?

- A. Pretraining the model on a specific domain
- B. Training the model using reinforcement learning
- C. Conditioning the model with task-specific instructions or demonstrations
- D. Adding more layers to the model

Answer: C

Explanation:

Comprehensive and Detailed In-Depth Explanation=

In-context learning is a capability of LLMs where the model adapts to a task by interpreting instructions or examples provided in the input prompt, without additional training. This leverages the model's pre-trained knowledge, making Option C correct. Option A refers to domain-specific pretraining, not in-context learning. Option B involves reinforcement learning, a different training paradigm. Option D pertains to architectural changes, not learning via context.

: OCI 2025 Generative AI documentation likely discusses in-context learning in sections on prompt-based customization.