Quiz - Generative AI Application Development

Generative AI Application Development

Back to course

Test passed Final score: 90 of 100 Test completed in: 29m 44s

It relies exclusively on a single Foundation Model to handle all reasoning and retrieval tasks to minimize latency.

It forces all data processing to occur within the GPU memory of the LLM to ensure data privacy.

It eliminates the need for prompt engineering by automating the intent classification process.

Correct answer Score: 5

Correct answer:
It tackles AI tasks using multiple interacting components, such as retrievers and external tools, allowing for more complex and reliable workflows.

Question 20 of 20

What is the primary benefit of using MLflow Tracing over standard logging when debugging Generative AI applications?

It compresses the logs into a binary format to save storage costs.

It automatically rewrites the Python code to optimize the latency of the chain.

It encrypts the logs so that only the model owner can view the prompts.

It allows for the interactive visualization of the call stack, tracking inputs, outputs, and latency for each span (step) of the chain.

Correct answer Score: 5

Quiz - Generative AI Application Development

Generative AI Application Development

Test passed Final score: 90 of 100 Test completed in: 29m 44s

It relies exclusively on a single Foundation Model to handle all reasoning and retrieval tasks to minimize latency.

It forces all data processing to occur within the GPU memory of the LLM to ensure data privacy.

It eliminates the need for prompt engineering by automating the intent classification process.

Correct answer Score: 5

Correct answer:
It tackles AI tasks using multiple interacting components, such as retrievers and external tools, allowing for more complex and reliable workflows.

Question 20 of 20

What is the primary benefit of using MLflow Tracing over standard logging when debugging Generative AI applications?

It compresses the logs into a binary format to save storage costs.

It automatically rewrites the Python code to optimize the latency of the chain.

It encrypts the logs so that only the model owner can view the prompts.

It allows for the interactive visualization of the call stack, tracking inputs, outputs, and latency for each span (step) of the chain.

Correct answer Score: 5