KillTest

Mejor calidad Mejor servicio

Examen

http://www.killtest.es

Renovación gratuita dentro de un año

Exam : **HPE0-V30**

Title : HPE AI Fundamentals

Version : DEMO

1.An AI Solutions Architect is evaluating models for a legal firm. The requirement is to analyze 15,000-word contracts and accurately link a definition on page 1 with a liability clause on page 40.
The architect rejects a legacy Long Short-Term Memory (LSTM) sequence-to-sequence model in favor of a modern Transformer architecture.
```
Project Constraints:
- Input Length: ~15,000 words per document.
- Accuracy Requirement: Exact linkage of distant entities.
- Hardware: NVIDIA DGX Cluster (A100 GPUs).
- Legacy System: LSTM with Bahdanau attention.
```
Why does the physical structure of the chosen Transformer guarantee superior accuracy for this specific long-document use case compared to the legacy LSTM?
A. The LSTM actively deletes its internal memory every 1,000 words to prevent GPU memory overflow, which inherently destroys the required cross-page linkages.
B. The Transformer utilizes a bidirectional recurrent loop that processes the document from back-to-front, capturing the liability clauses before the definitions.
C. The Transformer's self-attention computes a direct O(1) connection between any two words, eliminating sequential information decay and preserving long-range dependencies across the full document.
D. In legacy Transformer implementations with fixed context windows (e.g., BERT constrained to 512 tokens), documents are truncated into non-overlapping chunks. This avoids context confusion but explicitly prevents cross-page entity linkage required for legal analysis.
Answer: C

2.Which statement best describes the primary objective of cross-modal representation learning in the context of foundation models?
A. To compress high-resolution image files into sparse matrices using quantization techniques, with the primary goal of optimizing storage efficiency in vector database systems.
B. To strictly isolate audio, video, and text processing into completely independent neural network architectures to prevent data leakage during inference.
C. To directly convert raw text strings into pixel arrays through an end-to-end transformation process, explicitly avoiding any intermediate numerical vector representations or embedding layers.
D. To project data from fundamentally different modalities into a shared mathematical vector space for direct semantic similarity measurement.
Answer: D

3.Which statement correctly distinguishes the fundamental operational difference between a Chain and an Agent within the LangChain orchestration framework?
A. In its base implementation, an Agent is inherently stateless and cannot retain memory across user interactions, whereas a Chain automatically persists and manages session history in a robust backend database for subsequent requests.
B. A Chain dynamically selects and invokes external APIs based on real-time user intent analysis, whereas an Agent executes a predetermined Directed Acyclic Graph (DAG) structure for data ingestion

pipelines.
C. A Chain is strictly used for vector database indexing tasks, such as in RAG applications, whereas an Agent is exclusively responsible for processing the final natural language output presented to the end user.
D. An Agent leverages a language model to dynamically determine the next action, whereas a Chain follows a fixed, predetermined sequence of steps.
Answer: D

4.A Model Operations Analyst is reviewing a post-incident report regarding a multi-agent data processing pipeline. The system uses a "Map-Reduce" multi-agent design pattern to summarize 1,000 feedback surveys.
A "Mapper" agent is spawned concurrently 1,000 times to summarize each individual survey, and a single "Reducer" agent aggregates those 1,000 summaries into a final report.
```
==========================================================
Map-Reduce Agent Pattern Execution Audit
==========================================================
Total Input Surveys:          1000
Mapper Agents Spawned:        1000 (Concurrent)
Average Mapper Latency:       4.2 seconds
Reducer Agent Invocation:     Triggered at T+5.1 seconds
Final Output Status:          FAILED (HTTP 413 Payload Too Large)
Reducer Prompt Size:          185,000 tokens
==========================================================
```
Which TWO of the following represent architectural anti-patterns and flaws in this specific multi-agent implementation? (Choose 2.)
A. A Map-Reduce pattern is strictly incompatible with textual summarization tasks and should only be used for numerical aggregation in relational databases.
B. The Mapper agents were spawned concurrently instead of sequentially, preventing the system from utilizing the ConversationBufferMemory to pass state between surveys.
C. The Average Mapper Latency is too low, indicating that the Mapper agents bypassed the required vector database similarity search.
D. The Reducer agent is attempting to ingest all 1,000 Mapper summaries in a single monolithic prompt, violating the context window limits of standard generative models.
E. The system lacks an intermediate hierarchical aggregation layer (e.g., intermediate Reducers grouping 100 summaries at a time) to compress the payload before the final Reducer stage.
Answer: D, E

5.A DevOps Engineer is troubleshooting a newly deployed customer service multi-agent system built with LangGraph. The system consists of a "Greeting Agent," a "Technical Agent," and a "Billing Agent." The application frequently crashes with an out-of-memory (OOM) error or a max-token limit exception after several minutes of processing a single user query.
The engineer captures the following execution trace:

```
[00:01] Greeting_Agent: "I see you have a router issue. Routing to Technical."
[00:03] Technical_Agent: "I need to verify your account status first. Routing to Billing."
[00:06] Billing_Agent: "Account is active.
How can I help with your network?"
[00:08] Greeting_Agent: "I see you have a network issue. Routing to Technical."
[00:11] Technical_Agent: "I need to verify your account status first. Routing to Billing."
... (Pattern repeats continuously) ...
[05:42] SYSTEM ERROR: Token limit exceeded. Context window full.
```

Based on this diagnostic trace, which TWO of the following architectural flaws in the multi-agent orchestration design are causing this failure? (Choose 2.)

A. The Greeting_Agent was initialized with a temperature of 1.0, causing it to hallucinate the initial routing decision.

B. The agents are utilizing a shared global memory state where the historical intent is being overwritten, causing them to lose track of previously completed steps.

C. The system is deployed on a Kubernetes node without sufficient GPU resources, forcing the LLM to fallback to cyclic CPU processing.

D. The multi-agent graph lacks an explicit terminal node (END) or conditional exit logic, preventing the execution loop from concluding.

E. The vector database used for context retrieval has a stale index, providing outdated troubleshooting steps to the Technical Agent.

Answer: B, D