

1z0-1127-24 Exam Question & Interactive 1z0-1127-24 Testing Engine

1z0-1127-24 Exam Questions

Wednesday, July 10, 2024 8:56 AM

1. How are documents usually evaluated in the simplest form of keyword-based search?
 - a. According to the length of the documents
 - b. Based on the number of images and videos contained in the documents
 - c. By the complexity of language used in the documents
 - d. Based on the presence and frequency of the user-provided keywords

Google note: These indexing systems model the documents as a vector, i.e., a flat list of keywords. Keyword-based search is then achieved by looking at the relative frequencies of the various keywords in the documents and comparing them with the keywords in the user query, which is also modelled as a vector.

2. When is fine-tuning an appropriate method for customizing a Large Language Model (LLM)?
 - a. When you want to optimize the model without any instructions
 - b. When the LLM requires access to the latest data for generating outputs
 - c. When the LLM does not perform well on a task and the data for prompt engineering is too large
 - d. When the LLM already understands the topics necessary for text generation

Google note: Full finetuning is applicable to tasks where high accuracy is critical, we have access to a large amount of labeled data specifically customized to the target task and situations where the complexity of the task demands the full adaptability of the LLM architecture.

3. In which scenario is soft prompting appropriate compared to other training styles?
 - a. When the model requires continued pretraining on unlabeled data
 - b. When the model needs to be adapted to perform well in a domain on which it was not originally trained
 - c. When there is a need to add learnable parameters to a Large Language Model (LLM) without task-specific training
 - d. When there is a significant amount of labeled, task-specific data available

Google note: Soft prompts are crucial in ensuring precision and accuracy in AI-generated outputs by guiding a model's response in a specific direction. Unlike so-called hard prompts, which strictly define the input context, soft prompts provide general guidance while giving a model some flexibility in interpretation.

4. How does the temperature setting in a decoding algorithm influence the probability distribution over the vocabulary?
 - a. Increasing the temperature removes the impact of the most likely word.
 - b. Decreasing the temperature broadens the distribution, making less likely words more probable.
 - c. Increasing the temperature flattens the distribution, allowing for more varied word choices.
 - d. Temperature has no effect on probability distribution; it only changes the speed of decoding.

5. Which statement is true about Fine-tuning and Parameter-Efficient Fine-Tuning (PEFT)?
 - a. PEFT requires replacing the entire model architecture with a new one designed specifically for the new task, making it significantly more data-intensive than Fine-tuning.
 - b. Fine-tuning requires training the entire model on new data, often leading to substantial computational costs, whereas PEFT involves updating only a small subset of parameters, minimizing computational requirements and data needs.
 - c. Both Fine-tuning and PEFT require the model to be trained from scratch on new data, making them equally data and computationally intensive.
 - d. Fine-tuning and PEFT do not involve model modification; they differ only in the type of data used for training, with Fine-tuning requiring labeled data and PEFT using unlabeled data.

Google note: Better performance in low-data regimes: PEFT approaches have been shown to perform better than full fine-tuning in low-data regimes and generalize better to out-of-domain scenarios. Portability: PEFT methods enable users to obtain tiny checkpoints worth a few MBs compared to the large checkpoints of full fine-tuning.

AI Certification Page 1

P.S. Free & New 1z0-1127-24 dumps are available on Google Drive shared by Exam4Labs: <https://drive.google.com/open?id=1BfnRoGU7F2mRNa1hyCqL-0gJAIWX1ioN>

Our passing rate is very high to reach 99% and our 1z0-1127-24 exam torrent also boost high hit rate. Our 1z0-1127-24 study questions are compiled by authorized experts and approved by professionals with years of experiences. They are compiled according to the latest development conditions in the theory and practice and the questions and answers are based on real exam. Our study materials can improve your confidence for real exam and will help you remember the exam questions and answers that you will take part in. You can choose the version which suits you mostly. Our Oracle Cloud Infrastructure 2024 Generative AI Professional exam torrents simplify the important information and seize the focus to make you master the 1z0-1127-24 Test Torrent in a short time.

Before you try to attend the 1z0-1127-24 practice exam, you need to look for best learning materials to easily understand the key points of 1z0-1127-24 exam prep. There are 1z0-1127-24 real questions available for our candidates with accurate answers and detailed explanations. We are ready to show you the most reliable 1z0-1127-24 PDF VCE and the current exam information for your preparation of the test.

>> 1z0-1127-24 Exam Question <<

Hot 1z0-1127-24 Exam Question Pass Certify | Reliable Interactive 1z0-1127-24 Testing Engine: Oracle Cloud Infrastructure 2024 Generative AI

Professional

As long as you face problems with the 1z0-1127-24 exam, our company is confident to help you solve. Give our 1z0-1127-24 practice quiz a chance to give you a chance to succeed. We are very willing to go hand in hand with you on the way to preparing for 1z0-1127-24 Exam. And we have three different versions of our 1z0-1127-24 learning materials, you will find that it is so interesting and funny to study with our study guide.

Oracle 1z0-1127-24 Exam Syllabus Topics:

Topic	Details
Topic 1	<ul style="list-style-type: none">Fundamentals of Large Language Models (LLMs): For AI developers and Cloud Architects, this topic discusses LLM architectures and LLM fine-tuning. Additionally, it focuses on prompts for LLMs and fundamentals of code models.
Topic 2	<ul style="list-style-type: none">Using OCI Generative AI Service: For AI Specialists, this section covers dedicated AI clusters for fine-tuning and inference. The topic also focuses on the fundamentals of OCI Generative AI service, foundational models for Generation, Summarization, and Embedding.
Topic 3	<ul style="list-style-type: none">Building an LLM Application with OCI Generative AI Service: For AI Engineers, this section covers Retrieval Augmented Generation (RAG) concepts, vector database concepts, and semantic search concepts. It also focuses on deploying an LLM, tracing and evaluating an LLM, and building an LLM application with RAG and LangChain.

Oracle Cloud Infrastructure 2024 Generative AI Professional Sample Questions (Q32-Q37):

NEW QUESTION # 32

In LangChain, which retriever search type is used to balance between relevancy and diversity?

- A. similarity
- B. mmr
- C. top k
- D. similarity_score_threshold

Answer: B

Explanation:

In LangChain, the "mmr" (Maximal Marginal Relevance) search type is used to balance between relevancy and diversity when retrieving documents. This technique aims to select documents that are not only relevant to the query but also diverse from each other. This helps in avoiding redundancy and ensures that the retrieved set of documents covers a broader aspect of the topic. Maximal Marginal Relevance (MMR) works by iteratively selecting documents that have high relevance to the query but low similarity to the documents already selected. This ensures that each new document adds new information and perspectives, rather than repeating what is already included.

Reference

LangChain documentation on retrievers and search types

Research papers and articles on Maximal Marginal Relevance (MMR)

NEW QUESTION # 33

Which is NOT a built-in memory type in LangChain?

- A. Conversation ImgeMemory
- B. Conversation Buffer Memory
- C. Conversation Summary Memory
- D. Conversation Token Buffer Memory

Answer: A

NEW QUESTION # 34

Given a block of code:

qa = Conversational Retrieval Chain, from l1m (l1m, retriever-rtv, memory-memory) when does a chain typically interact with memory during execution?

- A. Before user input and after chain execution
- B. Continuously throughout the entire chain execution process
- C. Only after the output has been generated
- D. After user input but before chain execution, and again after core logic but before output

Answer: C

NEW QUESTION # 35

Which component of Retrieval-Augmented Generation (RAG) evaluates and prioritizes the information retrieved by the retrieval system?

- A. Generator
- B. Encoder-decoder
- C. Retriever
- D. Ranker

Answer: D

Explanation:

In Retrieval-Augmented Generation (RAG), the component responsible for evaluating and prioritizing the information retrieved by the retrieval system is the Ranker. After the Retriever fetches relevant documents or passages, the Ranker assesses these retrieved items based on their relevance to the query. It then prioritizes them, typically scoring and ordering the documents so that the most pertinent information is considered first in the generation process. This ensures that the generated response is based on the most relevant and useful content available.

Reference

Research papers on RAG (Retrieval-Augmented Generation)

Technical documentation on the architecture of RAG models

NEW QUESTION # 36

Which is a key characteristic of the annotation process used in T-Few fine-tuning?

- A. T-Few fine-tuning relies on unsupervised learning techniques for annotation.
- B. T-Few fine-tuning uses annotated data to adjust a fraction of model weights.
- C. T-Few fine-tuning involves updating the weights of all layers in the model.
- D. T-Few fine-tuning requires manual annotation of input-output pairs.

Answer: B

Explanation:

T-Few fine-tuning is a technique that uses annotated data to adjust only a fraction of the model's weights. This method aims to efficiently fine-tune the model with a limited amount of data and computational resources. By updating only a small subset of the parameters, T-Few fine-tuning can achieve significant performance improvements without the need for extensive training data or computational power.

Reference

Research papers on parameter-efficient fine-tuning techniques

Technical guides on T-Few fine-tuning methodology

NEW QUESTION # 37

.....

The Oracle desktop practice test software and web-based Understanding Oracle Cloud Infrastructure 2024 Generative AI

Interactive 1z0-1127-24 Testing Engine: <https://www.exam4labs.com/1z0-1127-24-practice-torrent.html>

- BTW, DOWNLOAD part of Exam4Labs 1z0-1127-24 dumps from Cloud Storage: <https://drive.google.com/open?id=1BfnRoGU7F2mRNalhyCqL-0gIAIWx1ioN>