Google.Generative-AI-Leader.v2026-08-21.q35

Exam Code:	Generative-AI-Leader
Exam Name:	Google Cloud Certified - Generative AI Leader Exam
Certification Provider:	Google
Free Question Number:	35
Version:	v2026-08-21
# of views:	110
# of Questions views:	375
https://www.freecram.net/torrent/Google.Generative-AI-Leader.v2026-08-21.q35.html	

NEW QUESTION: 1
What is the definition of generative AI?
A. A type of artificial intelligence that enables a system to autonomously learn and improve using neural networks and deep learning.4
B. A type of artificial intelligence that can create new content and ideas, including text, images, music, and code.
C. A type of machine learning algorithm inspired by the human brain that is made up of interconnected nodes.
D. A type of predictive model that estimates a relationship by fitting a line to the observed data.
Answer: (SHOW ANSWER)
The defining characteristic of generative AI is its ability to create new, original content that resembles its training data. This includes various modalities like text, images, music, and code, rather than just classifying, predicting, or analyzing existing data.

NEW QUESTION: 2
A development team is configuring a generative AI model for a customer-facing application and wants to ensure the generated content is appropriate and harmless. What is the primary function of the safety settings parameter in a generative AI model?
A. To limit the maximum text length that the model generates by ensuring concise responses.
B. To determine the number of tokens the model can process at once by influencing the complexity and length of inputs and outputs.
C. To filter out potentially harmful or inappropriate content from the model ' s output based on the desired level of filtering.
D. To control the creativity and randomness of the model ' s output by adjusting the diversity of word choices.
Answer: (SHOW ANSWER)

Safety settings in generative AI models are specifically designed to prevent the generation of content that could be harmful, offensive, or inappropriate. This includes filtering for categories like hate speech, sexually explicit content, self-harm, and violence, based on predefined thresholds. Options A, B, and D refer to other parameters like max_output_tokens or temperature, which control output length, input/output processing, and creativity, respectively, not safety.

NEW QUESTION: 3
A company's large learning model (LLM) is producing hallucinations that are a result of the Knowledge cutoff. How does retrieval-augmented generation (RAG) overcome this limitation?
A. RAG fine-tunes the LLM on specific customer query patterns to improve the speed and efficiency of response generation.
B. RAG enhances the creative writing capabilities of the LLM to generate more engaging and informative responses.
C. RAG enables the LLM to retrieve relevant and up-to-date information from knowledge sources.
D. RAG uses human oversight to ensure accuracy before presenting information to the customer.
Answer: (SHOW ANSWER)
The primary purpose of RAG is to address the " knowledge cutoff " and hallucination issues of LLMs. It does this by retrieving relevant, up-to-date information from external knowledge sources (like databases or documents) at inference time and then using this retrieved information to ground the LLM ' s generation, ensuring factual accuracy and relevance to the specific query.

NEW QUESTION: 4
A global news company is using a large language model to automatically generate summaries of news articles for their website. The model ' s summary of an international summit was accurate until it hallucinated by stating a detail that did not occur. How should the company overcome this hallucination?
A. Implement stricter safety settings to filter out potentially controversial topics.
B. Fine-tune the model on a larger dataset of news articles.
C. Increase the temperature setting of the model to encourage more diverse outputs.
D. Use grounding to base the model output on the source articles.
Answer: (SHOW ANSWER)
The core problem is the model ' s hallucination-it invented a factual detail-in a context (news reporting) where factual accuracy is non-negotiable. To correct a factual error in a generative summary, the model must be constrained to speak only based on verifiable facts from a reliable source.
The most effective technique to combat hallucinations and ensure factual adherence is Grounding (D).
Grounding connects the Large Language Model ' s (LLM ' s) output to a specific, trusted, and verifiable source of information. This is often implemented using Retrieval-Augmented Generation

(RAG). In this scenario, grounding the summary model on the original source articles ensures that every generated statement is directly entailed by the provided facts (the source article content). Option B, fine-tuning, is expensive and only updates the model ' s general knowledge and style; it does not prevent the model from guessing or fabricating details when retrieving information. Option C, increasing temperature, would make the output less consistent and more diverse, likely increasing the chance of hallucination, which is the opposite of the desired effect. Option A is unrelated to factual accuracy. Therefore, Grounding is the necessary step to anchor the model ' s responses to the true content of the source articles.
(Reference: Google Cloud documentation on RAG/Grounding emphasizes that its primary purpose is to address the "knowledge cutoff" and hallucination issues of LLMs by retrieving relevant, up-to-date information from external knowledge sources and using this retrieved information to ground the LLM ' s generation, ensuring factual accuracy.)

NEW QUESTION: 5
What are core hardware components of the infrastructure layer in the generative AI landscape?
A. TPUs and GPUs
B. User interfaces
C. Pre-trained models
D. Tools and services for building AI models
Answer: (SHOW ANSWER)
The Generative AI landscape is often broken down into several functional layers: Applications, Agents, Platforms, Models, and Infrastructure.
The Infrastructure Layer is the foundation, providing the physical and virtual computing resources necessary to run and train the large models. These resources include servers, storage, networking, and most importantly, the specialized hardware accelerators required for high-volume, parallel computation.
The core hardware components are the Graphics Processing Units (GPUs) and the custom-designed Tensor Processing Units (TPUs) (A). These accelerators are optimized for the massive matrix operations fundamental to deep learning and Gen AI model training and inference.
Options B (User interfaces) and D (Tools and services) refer to the Application and Platform layers, respectively.
Option C (Pre-trained models) refers to the Model layer.
The physical hardware underpinning these abstract layers are the TPUs and GPUs.
(Reference: Google Cloud Generative AI Study Guides state that the Infrastructure Layer provides the core computing resources needed for generative AI, including the physical hardware (like servers, GPUs, and TPUs) and the essential software needed to train, store, and run AI models.)

NEW QUESTION: 6
A company is trying to decide which platform to use to optimize its generative AI (gen AI) solutions. Why should the company use Vertex AI Platform?

A. It provides a mechanism for efficient analysis and exploration of large datasets used in machine learning.
B. It provides gen AI coding assistance with enterprise security and privacy protection.
C. It provides scalable and cost-effective object storage for data used in machine learning workflows.
D. It provides a unified platform of tools for building, deploying, and managing machine learning.
Answer: (SHOW ANSWER)
Vertex AI is Google Cloud's core, end-to-end Machine Learning Operations (MLOps) platform, designed to cover the entire ML lifecycle.
The key benefit of Vertex AI, particularly for generative AI, is that it provides a unified platform (D) where all stages of AI development-from accessing foundation models in Model Garden, testing in Vertex AI Studio, training and tuning (via tools like Reinforcement Learning from Human Feedback), to deploying, and monitoring models in production-can be managed from a single service. This significantly reduces complexity, improves collaboration between teams (data scientists, engineers, business leaders), and ensures enterprise-grade governance and scalability necessary for production Gen AI solutions.
Option A describes BigQuery.
Option B describes Gemini Code Assist.
Option C describes Cloud Storage.
Vertex AI is the overarching platform that integrates all these tools to deliver a streamlined MLOps workflow.
(Reference: Google Cloud documentation states that Vertex AI is the unified AI development platform that brings together Google Cloud services for building, deploying, and managing machine learning models and generative AI solutions.)

NEW QUESTION: 7
A company is defining their generative AI strategy. They want to follow Google-recommended practices to increase their chances of success. Which strategy should they use?
A. Rapid implementation strategy
B. Bottom-up strategy
C. Multi-directional strategy
D. Top-down strategy
Answer: (SHOW ANSWER)
Google Cloud often recommends a "top-down" approach for generative AI strategy. This means starting with clear business objectives and leadership alignment on how generative AI can solve critical business problems, rather than simply experimenting from the bottom up without a clear strategic direction.

NEW QUESTION: 8

A user asks a generative AI model about the scientific accuracy of a popular science fiction movie. The model confidently states that humans can indeed travel faster than light, referencing specific but entirely fictional theories and providing made-up explanations of how this is achieved according to the movie's "established science." The model presents this information as factual, without indicating that it originates from a fictional work. What type of model limitation is this?

A. Bias

B. Knowledge cutoff

C. Data dependency

D. Hallucination

Answer: (SHOW ANSWER)

The limitation described is the AI model generating a false or misleading response (humans traveling faster than light is scientifically impossible/unproven) and presenting it as fact (confidently stating a fictional theory is real) without the ability to indicate its uncertainty or the source's fictional nature. This is the definition of a Hallucination in generative AI.

AI Hallucinations occur when a Large Language Model (LLM) generates outputs that are factually incorrect, irrelevant, or nonsensical, despite being linguistically fluent and seemingly plausible. They arise because the model is designed to predict the most statistically probable next word or token based on its training data, even when it lacks information or when its training data contains a mixture of fact and fiction. The model is overconfident in its generated response, a behavior that diminishes user trust and reliability, especially in applications where factual accuracy is critical.

While a knowledge cutoff (B) is a common cause of hallucinations when an LLM is asked about recent events, the core limitation of fabricating facts from its own hardwired knowledge is the hallucination itself. Data dependency (A) relates to the model's reliance on the quality and completeness of its training data, and while flawed training data can be a cause, the error mode of inventing facts is the Hallucination.

NEW QUESTION: 9

A large company is creating their generative AI (gen AI) solution by using Google Cloud ' s offerings. They want to ensure that their mid-level managers contribute to a successful gen AI rollout by following Google- recommended practices. What should the mid-level managers do?

A. Perform continuous testing, measurement, and refinement based on user feedback and real-world performance data.

B. Create a robust data strategy to ensure teams can access high-quality, relevant data that is appropriate for training and fine-tuning gen AI models.

C. Drive gen AI adoption by identifying high-impact, feasible solutions that address specific challenges within their workflows.

D. Secure funding and resources for AI initiatives by demonstrating the potential return on investment to the chief financial officer (CFO).

Answer: (SHOW ANSWER)

Google ' s recommended strategy for a successful generative AI rollout involves a combination of top-down strategic alignment and bottom-up adoption. In this structure, the role of the mid-level manager is critical for driving tangible value within their specific domain.
Securing funding (D) is typically the responsibility of senior leadership or the steering committee.
Creating a robust data strategy (B) is the domain of data governance teams and data scientists.
Continuous testing and refinement (A) is the job of MLOps/engineering teams and end-users.
The primary role of the mid-level manager is to act as the bridge between high-level strategy and daily operations. They possess the domain knowledge to pinpoint pain points. Therefore, their most impactful contribution is to identify specific, high-impact, and feasible use cases (C) for their teams-such as automating report summaries or drafting internal communications-that directly address operational challenges and demonstrate quick wins. This action fuels successful adoption and validates the AI strategy from the ground up.
(Reference: Google Cloud ' s guidance on Gen AI strategy emphasizes that successful adoption requires strong top-down vision (like defining goals/funding) combined with bottom-up discovery, where functional leaders (mid-level managers) identify and prioritize high-value, feasible solutions within their specific workflows to drive adoption.)

NEW QUESTION: 10

A home loan company is deploying a generative AI system to automate initial loan application reviews. Several applicants have been unexpectedly rejected, leading to customer complaints and potential bias concerns. They need to ensure responsible and fair lending practices. What aspect of the AI system should they prioritize?

A. Implementing stricter data security measures to protect applicants ' financial information from unauthorized access.
B. Ensuring AI decision-making is explainable to understand decision reasons and establish accountability.
C. Increasing the speed at which the AI system processes loan applications to handle the high volume.
D. Regularly updating the AI model with more financial data to improve its accuracy over time.
Answer: B (LEAVE A REPLY)
The problem centers on unexpected rejections and potential bias in a high-stakes, regulated domain (lending).
In such a context, the central tenet of Responsible AI is transparency and fairness.
While all options are valid goals, the priority when facing bias concerns and customer complaints due to rejection is to provide accountability and verify the fairness of the automated decision. This is achieved through Explainable AI (XAI).
Ensuring AI decision-making is explainable (B) means building mechanisms that allow developers, regulators, and affected customers to understand why a specific decision (rejection) was made. Explainability is crucial for:
Auditing for bias: If the reasons for rejection can be traced (e.g., system rejects based on loan-to-value ratio, not race), bias can be identified and corrected.

Compliance: Financial services are heavily regulated, and the ability to explain a lending decision is often a legal or regulatory requirement.
Customer Trust: Providing a clear reason for rejection (even if the news is bad) reduces complaints and fosters confidence, directly addressing the core issue of unexpected rejections.
Options A, C, and D address security, speed, and accuracy, respectively, but Explainability is the direct mechanism for proving fairness and ensuring accountability, making it the most critical priority in this scenario.
(Reference: Google ' s Responsible AI principles and training materials highlight that in high-stakes domains like finance, explainability is essential for establishing trust, identifying and mitigating bias, and meeting regulatory compliance.)

NEW QUESTION: 11
What will Google Cloud's Agent Assist help a company achieve?
A. The infrastructure to provide an enterprise-grade contact center solution with omnichannel support, routing, and integration with CRM systems.
B. The ability to analyze conversational data to identify customer sentiment, common topics of discussion, and insights into agent performance and customer experience.
C. The ability to provide real-time assistance and recommended responses to live customer service agents during their interactions.
D. The ability to build and deploy deterministic and generative chatbot agents for automated customer support.
Answer: (SHOW ANSWER)
Google Cloud's Agent Assist is specifically designed to augment human customer service agents. It provides real-time suggestions, retrieves relevant information, and offers recommended responses to agents during live interactions, improving their efficiency and consistency.

NEW QUESTION: 12
A learning and development team wants to quickly create a new hire training video with a custom avatar and voiceover that matches their company's branding and key messaging. They did not receive any money to spend on the production. What should they do?
A. Generate the video frames with Imagen.
B. Prompt the Gemini app to create a video.
C. Train a model with Vertex AI and produce a video.
D. Create a video with Google Vids.
Answer: (SHOW ANSWER)
The scenario requires quick creation of a training video using a custom avatar and voiceover while adhering to zero cost for production.
Google Vids is an AI-powered video creation app (part of Google Workspace/Gemini features) designed to make video creation accessible for teams without the overhead of traditional production. It specifically offers features like AI avatars and voiceovers for content such as

trainings, demos, and onboarding videos. This directly addresses the need for a low-cost, fast solution for a new hire training video with custom branding elements (custom avatars and voiceovers are a key feature of the tool).
Option A, Imagen, is a Google foundation model specialized for image generation, not the creation of structured, narrated training videos with avatars. Option B, using the Gemini app, is primarily for text, code, and multimodal chat/generation, and is not the dedicated Google application for video production. Option C, training a model with Vertex AI, is a highly technical, time-consuming, and expensive endeavor that violates the need for a quick and zero-cost solution. Therefore, using the purpose-built, gen AI-enabled Google Vids application is the correct and most efficient choice.

NEW QUESTION: 13
A company is exploring Gemini Enterprise (Agentspace) to improve how its employees search for information on their enterprise systems and automate certain tasks. What is the key business advantage of using Gemini Enterprise (Agentspace)?
A. More granular control over support team access and permissions for sensitive data.
B. Enhanced real-time communication and collaboration among team members.
C. Improved productivity and data interaction using AI assistants and advanced document analysis.
D. Greater interoperability with legacy software systems and databases.
Answer: (SHOW ANSWER)
Gemini Enterprise (Agentspace) is designed as an enterprise-grade AI environment built to solve information fragmentation and employee productivity issues.
The key business advantage of this platform is improved productivity and data interaction using AI assistants and advanced document analysis (C). Agentspace enables organizations to centralize access to internal knowledge bases, document repositories, and communication channels. Employees can use conversational AI assistants to immediately query vast libraries of unstructured corporate data, extract key performance metrics, summarize massive compliance documents, and execute workflow automations without leaving their primary working environment. This drastically reduces time spent manually tracking down information across fragmented tools.
* Option A relates to Identity and Access Management (IAM) or basic data governance controls, which are prerequisite security frameworks rather than the unique business value proposition of Agentspace.
* Option B describes a communication tool like Google Chat or Slack.
* Option D describes specialized middleware or enterprise service buses (ESB), whereas Agentspace focuses on intelligent interaction and synthesis layer rather than base database protocol interoperability.
(Reference: Google Cloud Workspace and Gemini Enterprise strategic whitepapers state that Agentspace serves as a centralized hub that transforms employee workflows by embedding

conversational AI assistants into corporate data repositories, unlocking advanced document analysis to maximize knowledge worker velocity and overall productivity.)

NEW QUESTION: 14
A pharmaceutical company's research and development department spends significant time manually reviewing new scientific papers to identify potential drug targets. They need a solution that can answer questions about these documents and provide summarized insights to researchers without requiring extensive coding expertise. What should the organization do?
A. Use Gemini for Google Workspace to facilitate collaborative document review.
B. Use Vertex AI Search to index the papers and enable keyword-based searches.
C. Use Vertex AI AutoML to train a model that classifies papers into predefined research areas.
D. Use Vertex AI Agent Builder to create a custom AI agent.
Answer: (SHOW ANSWER)
The requirement is to answer questions about the documents and provide summarized insights without requiring extensive coding expertise. Vertex AI Agent Builder is designed precisely for creating custom AI agents, often with low-code or no-code capabilities, that can interact with and process large volumes of information like scientific papers. While Vertex AI Search could index papers for keyword searches, it doesn't directly answer questions or provide summarized insights in the same way a generative AI agent built with Agent Builder could. Gemini for Google Workspace is for collaborative work, not specifically for building custom AI agents for document analysis. Vertex AI AutoML is for training classification models, which is different from answering questions and summarizing.

NEW QUESTION: 15
A human resources team is implementing a new generative AI application to assist the department in screening a large volume of job applications. They want to ensure fairness and build trust with potential candidates. What should the team prioritize?
A. Integrating the AI application with various job boards to maximize candidate reach.
B. Focusing on minimizing the processing time for each application to improve efficiency.
C. Ensuring AI operates transparently, especially regarding application evaluation and data usage.
D. Ensuring that the AI application can automatically rank all candidates without requiring human review.
Answer: (SHOW ANSWER)
To ensure fairness and build trust, especially in sensitive areas like job applications, transparency in how AI evaluates applications and uses data is paramount. This involves understanding potential biases, explaining decisions (where possible), and ensuring human oversight.

NEW QUESTION: 16

A development team is building an internal knowledge base chatbot to answer employee questions about company policies and procedures. This information is stored across various documents in Google Cloud Storage and is updated regularly by different departments. What is the primary benefit of using Google Cloud
' s RAG APIs in this scenario?
A. They provide a pre-built user interface for the chatbot, simplifying the front-end development process.
B. They allow the development team to train a single foundation model on all company documents.
C. They enable the generative AI model to retrieve the most up-to-date and relevant information from the policy documents in real-time.
D. They automatically create summaries of all company policies, which are then presented to employees as quick answers.
Answer: (SHOW ANSWER)
The primary benefit of RAG (Retrieval-Augmented Generation) in this context is its ability to ensure the chatbot provides accurate and up-to-date information. By retrieving relevant and recent policy documents from Cloud Storage in real-time and then grounding the LLM ' s response with this information, the chatbot avoids hallucinating or providing outdated answers, which is crucial for an internal knowledge base.

NEW QUESTION: 17
A company is developing a generative AI-powered customer support chatbot. They want to ensure the chatbot can answer a wide range of customer questions accurately, even those related to recently updated product information not present in the model ' s original training data. What is a key benefit of implementing retrieval- augmented generation (RAG) in this chatbot?
A. RAG will significantly reduce the computational resources required to run the generative AI model.
B. RAG will primarily help the chatbot generate more creative and engaging conversational responses.
C. RAG will enable the chatbot to fine-tune its underlying language model on the fly based on customer interactions.

D. RAG will enable the chatbot to access and utilize external, up-to-date knowledge sources to provide more accurate and relevant answers.

Answer: (SHOW ANSWER)

The central problem is the Large Language Model ' s (LLM ' s) knowledge cutoff, where it cannot answer questions about information that appeared after its training data was collected (e.g., recently updated product details).

Retrieval-Augmented Generation (RAG) is specifically designed to overcome this limitation. The process involves:

Retrieval: When a question is asked, the RAG system first searches an external, up-to-date knowledge source (like a vector database of current product docs).

Augmentation: It retrieves the most relevant, recent text snippets (the context).

Generation: This retrieved context is added to the user ' s prompt (augmentation) and sent to the LLM, forcing the model to ground its response in the current facts.

The key benefit is thus to enable the chatbot to access and utilize external, up-to-date knowledge sources (D).

This ensures the answers are accurate and relevant to the most current product information, directly addressing the knowledge cutoff issue without requiring expensive model retraining.

Option B is the function of the Temperature setting, not RAG.

Option C describes an unproven and unscalable model update mechanism (fine-tuning is a separate process).

RAG is a process enhancement that prioritizes accuracy and relevance over merely reducing computation (A).

(Reference: Google Cloud documentation on RAG states that its primary purpose is to address the

"knowledge cutoff" and hallucination issues of LLMs by retrieving relevant and up-to-date information from external knowledge sources at inference time and using this retrieved information to ground the LLM ' s generation, ensuring factual accuracy.)

NEW QUESTION: 18

A research company needs to analyze several lengthy PDF documents containing financial reports and identify key performance indicators (KPIs) and their trends over the past year. They want a Google Cloud prebuilt generative AI tool that can process these documents and provide summarized insights directly from the source material with citations. What should the analyst do?

A. Create a custom Gem in Gemini Advanced with predefined KPIs to look across different financial reports.

B. Use the Gemini app to ask general financial trend questions.

C. Use NotebookLM to upload and analyze the documents.

D. Use Gemini for Google Workspace within Google Docs to copy and paste sections of the reports for summary and analysis.

Answer: (SHOW ANSWER)

The requirements are for a prebuilt tool that is designed for:

Analyzing uploaded private documents (lengthy PDFs).
Providing summarized insights (extracting KPIs and trends).
Offering citations (grounding the answers to the source material).
NotebookLM (C) is the Google tool explicitly designed for this use case. It is a generative AI powered notebook/research assistant that allows users to upload source documents (including PDFs), then ask questions and generate summaries or insights that are grounded in and cited back to the source documents.
This makes it an ideal prebuilt solution for an analyst who needs to process complex, lengthy financial reports and verify the data with citations.
Gemini Advanced (A) and Gemini app (B) are general-purpose conversational tools that are not primarily focused on deep, grounded analysis of uploaded documents that require source citations for research integrity.
Gemini for Google Workspace (D) is limited to data already in Workspace apps (Docs, Gmail, Drive) and the manual copy/paste process would be inefficient for " several lengthy PDF documents. " (Reference: Google ' s Generative AI Leader training materials highlight NotebookLM as the specific generative AI application built for research and information synthesis from uploaded documents, offering key features like grounding and citations back to the source material.)

NEW QUESTION: 19
A company wants to use an AI agent to automate some tasks. They want everyone to understand the different functions of an AI agent. What is the function of an AI agent in the context of gen AI?
A. To provide the computational resources needed to train and run gen AI models.
B. To store and manage large datasets used for training and running gen AI models.
C. To provide a user-friendly interface for interacting with gen AI models.
D. To analyze situations, use multiple tools, and make informed decisions without requiring constant human input.
Answer: (SHOW ANSWER)
An AI agent, especially in the context of generative AI, is designed to be more autonomous and capable than a simple model. Its function is to understand a goal, analyze a situation, leverage various tools (including other generative AI models or external APIs), and make decisions or take actions to achieve that goal, often with minimal human intervention.

NEW QUESTION: 20
A company is developing a conversational AI chatbot. They need to ensure the chatbot can engage in human- like conversations and provide accurate information. What should they do to enhance thechatbot's ability to understand and respond effectively to user prompts?

A. Use prompt engineering techniques, like few-shot prompting, to provide the chatbot with examples of successful interactions.
B. Limit the chatbot's training data to prevent it from learning irrelevant information.

C. Use strict keyword matching to ensure that the chatbot only responds to specific commands.
D. Lower model temperature setting to produce more consistent and predictable responses.
Answer: (SHOW ANSWER)
Prompt engineering, especially techniques like few-shot prompting (providing examples of desired input- output pairs), is crucial for guiding a generative AI model to understand context and generate relevant, human- like responses. Limiting data or using strict keyword matching would severely restrict the chatbot's conversational ability, and lowering temperature makes responses less creative, not necessarily more understanding.

NEW QUESTION: 21
A financial services company receives a high volume of loan applications daily submitted as scanned documents and PDFs with varying layouts. The manual process of extracting key information is time- consuming and prone to errors. This causes delays in loan processing and impacts customer satisfaction. The company wants to automate the extraction of this critical data to improve efficiency and accuracy. Which Google Cloud tool should they use?
A. Natural Language API
B. Dataflow
C. Vision AI
D. Document AI API
Answer: (SHOW ANSWER)
Document AI API is specifically designed for intelligent document processing. It uses machine learning to extract structured data from unstructured documents like scanned forms and PDFs, even with varying layouts.
This directly addresses the challenge of automating data extraction from loan applications. Natural Language API focuses on text understanding, Vision AI on image analysis (not structured extraction from documents), and Dataflow is for data processing pipelines.

NEW QUESTION: 22
A company wants to use generative AI to create a chatbot that can answer customer questions about their products and services. They need to ensure that the chatbot only uses information from the company ' s official documentation. What should the company do?
A. Use role prompting.
B. Adjust the temperature parameter.
C. Use prompt chaining.
D. Use grounding.
Answer: (SHOW ANSWER)
The core requirement is to guarantee that the chatbot only uses information from the company ' s official documentation and does not rely on its general knowledge base. This is crucial for

ensuring factual accuracy, relevance to the company ' s specific products, and preventing the generation of fabricated or incorrect information (hallucinations).
The specific technique designed to address this challenge is Grounding. Grounding is the process of connecting the Large Language Model ' s (LLM ' s) responses to a trusted, verifiable source of information, such as an organization ' s internal documents, databases, or live data feeds. When an LLM is grounded, it is forced to base its answers only on the provided context, effectively preventing it from drawing on its broad, generalized training data. Grounding is often implemented using a method called Retrieval-Augmented Generation (RAG), particularly with tools like Google Cloud ' s Vertex AI Search, which indexes the official documentation and feeds the relevant snippets to the model.
Options A, B, and C address different aspects of model output: Role prompting sets the model ' s persona, adjusting temperature controls creativity, and prompt chaining manages conversation history, but none of these techniques restrict the model ' s source of truth to the official documentation. Therefore, Grounding is the correct and most effective technique for this requirement.

NEW QUESTION: 23
A security team needs a centralized platform to gain a comprehensive overview of their organization's security health across their entire Google Cloud environment, including potential threats to their generative AI deployments. Which Google Cloud security offering is specifically for this purpose?
A. Workload monitoring tools
B. Security Command Center
C. Identity and Access Management
D. Secure-by-design infrastructure
Answer: (SHOW ANSWER)
Security Command Center is Google Cloud's comprehensive security management and data risk platform. It provides centralized visibility into security posture, identifies vulnerabilities, detects threats, and helps manage compliance across the entire Google Cloud environment, includingservices and deployments like generative AI.

NEW QUESTION: 24
A finance team wants to use Gemma to help with daily tasks so that the financial analysts can focus on other work. Which business problem can Gemma most efficiently address?
A. The complexity of building and deploying sophisticated internal knowledge bases to answer employees
' finance-related questions with accurate and up-to-date information.
B. The difficulty in analyzing large datasets of financial transactions and market data to identify anomalies and predict future financial performance.

C. The struggle to accurately extract key financial figures and insights from a variety of document formats, such as balance sheets and income statements, for quick reporting.
D. The challenge of efficiently producing high-quality written summaries and initial drafts of financial communications.
Answer: (SHOW ANSWER)
Gemma is a family of lightweight, open-source Large Language Models (LLMs) from Google that are based on the same research and technology as the Gemini models. As an LLM, its core strength lies in language- based tasks, particularly the generation and summarization of text.
The problem that Gemma, or any pure LLM, can most efficiently address is:
Generating text: creating new content quickly (Option D).
Summarizing text: condensing long communications or documents (Option D).
Option D, producing high-quality written summaries and initial drafts, is a natural language generation task that aligns perfectly with the core function of an LLM like Gemma. It is a key productivity booster for analysts needing to draft reports or emails quickly.
Option B (Analyzing large datasets/predicting performance) requires traditional machine learning (ML) models or analytical tools like BigQuery ML, as LLMs are not specialized for numerical predictive modeling.
Option C (Extracting key financial figures from documents) is a task for a highly specialized tool like Google
' s Document AI.
Option A (Building internal knowledge bases for Q & A) is a broader use case that is best solved with a platform solution using RAG, such as Vertex AI Search, not just a base model.
(Reference: Google ' s description of the Gemma model family emphasizes its role as a flexible, open LLM that excels at language fundamentals, making it ideal for content creation, summarization, and other text generation tasks.)

NEW QUESTION: 25
A retail company with a large online catalog wants to improve customer experience and drive sales by implementing multimodal search capabilities (image, voice, and text). What is a primary business benefit of this capability?
A. Improved customer engagement and product discovery leading to increased satisfaction and potential sales.
B. Reduced dependency on keyword optimization for product listings and improved search engine rankings.
C. Lowered operational costs associated with managing and updating product information across different platforms and channels.
D. Streamlined inventory management processes and more accurate demand forecasting for popular items.
Answer: (SHOW ANSWER)
Multimodal search directly enhances the customer experience by allowing them to find products using various intuitive methods (images, voice, text). This leads to easier product discovery,

higher engagement, and ultimately increased customer satisfaction and potential sales, which is a primary business benefit.

NEW QUESTION: 26
An organization wants to use generative AI to create a marketing campaign. They need to ensure that the AI model generates text that is appropriate for the target audience. What should the organization do?
A. Use role prompting.
B. Use prompt chaining.
C. Use few-shot prompting.
D. Adjust the temperature parameter.
Answer: (SHOW ANSWER)
Role prompting is a technique where you instruct the generative AI model to "act as" a specific persona or character. By assigning the model a role (e.g., "Act as a marketing expert writing for a young, tech-savvy audience"), you can guide its tone, style, and content to be appropriate for the target audience of the marketing campaign.

NEW QUESTION: 27
A marketing team wants to use a generative AI model to create product descriptions for their new line of eco-friendly water bottles. They provide a brief prompt stating, "Write a product description for our new water bottle." The model generates a generic, lackluster description that is factually accurate but lacks engaging language and doesn't highlight the environmental benefits that are key to their brand. What should the marketing team do to overcome this limitation of the generated product description?
A. Train the model on a dataset of marketing materials from other eco-friendly brands.
B. Add details to the prompt about the audience, tone, and keywords.
C. Increase the token count for the model to allow for longer descriptions.
D. Lower the temperature setting of the model to produce more consistent results.
Answer: (SHOW ANSWER)
The core problem described is a lackluster and generic output that fails to capture the desired tone and key information (environmental benefits). This is a classic limitation of zero-shot prompting (a brief, un-detailed prompt), where the generative AI model relies solely on its general training data and lacks the necessary context to produce a highly relevant and engaging response. The solution is to improve the quality of the prompt itself, a process known as Prompt Engineering.
Option A, training the model, is an expensive and time-consuming process (fine-tuning) that is usually unnecessary for stylistic or content-specific guidance that can be achieved with a good prompt. Options C and D control the length and creativity, respectively, but don't inject the missing information or brand requirements.

Adding details to the prompt is the most immediate and effective technique to guide the model. By specifying the target audience (e.g., eco-conscious consumers), the desired tone (e.g., enthusiastic, persuasive), and mandatory keywords (e.g., "sustainable," "BPA-free," "ocean-friendly"), the marketing team is effectively providing the model with the necessary constraints and context to produce a description that is tailored to their brand and marketing goals. This technique is fundamental to improving the output of generative AI models without resorting to model customization.

NEW QUESTION: 28
An organization wants granular control over who can use and see their generative AI models and related resources on Google Cloud. Which Google Cloud security offering is specifically for this purpose?
A. Identity and Access Management
B. Secure-by-design infrastructure
C. Security Command Center
D. Workload monitoring tools
Answer: A (LEAVE A REPLY)
Identity and Access Management (IAM) is the fundamental Google Cloud service that allows you to define who has what access to which resources. It provides granular control over permissions for users, groups, and service accounts, including access to generative AI models and related data.

NEW QUESTION: 29
A company is exploring Google Agentspace to improve how its employees search for information on their enterprise systems and automate certain tasks. What is the key business advantage of using Agentspace?
A. Enhanced real-time communication and collaboration among team members.
B. Greater interoperability with legacy software systems and databases.
C. Improved productivity and data interaction using AI assistants and advanced document analysis.
D. More granular control over support team access and permissions for sensitive data.
Answer: C (LEAVE A REPLY)
Google Agentspace (or similar agent platforms) is designed to empower employees with AI-powered assistants that can navigate and interact with enterprise systems, analyze documents, and automate tasks. This directly leads to improved employee productivity and more efficient data interaction by leveraging AI to streamline workflows and provide faster access to information.

NEW QUESTION: 30

A team is using a generative AI model to automatically generate short summaries of customer feedback. They need to ensure that these summaries are concise and easy to digest. What model setting should they adjust?

A. Top-p (nucleus sampling)
B. Safety settings
C. Temperature
D. Output length
Answer: (SHOW ANSWER)
The objective is to make the generated summaries concise-that is, to control their length.
In the configuration of a generative AI model, particularly a large language model (LLM), the parameter used to directly control the maximum size of the response is the Output Length parameter (often referred to as max_output_tokens or max_tokens). By setting a low limit on this parameter, the team can ensure that the model is forced to terminate its response once that limit is reached, resulting in a shorter, more concise summary that is " easy to digest, " as requested.
The other parameters control different aspects of the output quality:
Temperature (C) controls the creativity or randomness of the output. Lowering it makes the output more predictable; raising it makes it more diverse. It does not control length.
Top-p (A) is a decoding method related to temperature that also controls the model ' s creativity by limiting the vocabulary from which it can choose the next token. It does not control length.
Safety settings (B) are used to filter and block the generation of harmful, illegal, or inappropriate content.
They do not affect the length or conciseness of the output.
(Reference: Google Cloud ' s Generative AI documentation on model parameters explicitly lists max_output_tokens or Output Length as the setting used to determine the maximum size of a model ' s generated response.)

NEW QUESTION: 31
What is a characteristic of Google Cloud as a generative AI company?
A. Google Cloud has an AI-first focus that enables innovation, with continuous updates and broad integration across its platform.
B. Google Cloud ensures that all generative AI models and data are completely secured and isolated from external networks.
C. Google Cloud provides fully autonomous AI agents that require zero configuration or management overhead.
D. Google Cloud relies on proprietary, closed-source AI technologies for maximum security benefits.
Answer: (SHOW ANSWER)

NEW QUESTION: 32
An organization wants to use generative AI to create a chatbot that can answer customer questions about their account balances. They need to ensure that the chatbot can access previous portions of the conversation with the customer. Which prompting technique should they use?
A. Use zero-shot prompting.
B. Use role prompting.
C. Use few-shot prompting.
D. Use prompt chaining.
Answer: (SHOW ANSWER)
Prompt chaining (or conversational memory/context management) is the technique used to maintain the conversational context. It involves feeding previous turns of a conversation (or a summary of them) back into the model along with the current user query, allowing the chatbot to "remember" and reference past interactions for coherent and contextually relevant responses, especially crucial for tasks like checking account balances that span multiple turns.

NEW QUESTION: 33
A marketing team wants to use a foundation model to create social media and advertising campaigns. They want to create written articles and images from text. They lack deep AI expertiseand need a versatile solution.
Which Google foundation model should they use?
A. Gemma
B. Imagen
C. Gemini
D. Veo
Answer: (SHOW ANSWER)
Gemini is Google's most advanced and multimodal foundation model, capable of understanding and generating various forms of content, including text and images, from a single prompt. Its versatility makes it suitable for marketing teams that need to create diverse campaign materials without deep AI expertise.
Imagen is specifically for image generation, Gemma is a family of smaller, open models, and Veo is for video generation.

NEW QUESTION: 34
What does a diffusion model do?
A. Analyzes data and predicts future trends and patterns.
B. Optimizes business processes and resource allocation.
C. Facilitates the storage and management of structured data.
D. Generates high-quality content by refining noise into structured data.
Answer: (SHOW ANSWER)
A Diffusion Model (or Denoising Diffusion Probabilistic Model) is a specific class of generative AI model that is best known for its ability to create highly realistic images (e.g., Google's Imagen and Stable Diffusion are based on this architecture).
The core mechanism of a diffusion model is a two-step process:
Forward Diffusion (Adding Noise): It learns how to gradually corrupt data (like an image) by adding random noise until the original content is completely indistinguishable.
Reverse Diffusion (Denoising): It then learns to reverse this process-to gradually remove the noise-starting from a random noise pattern and iteratively refining it, guided by a text prompt, until a clear, coherent, and high-quality piece of content (an image or video) is generated.
Option D accurately captures this mechanism: the model starts with pure noise and generates the final structured data (the image) by refining that noise.
Option A describes predictive AI (forecasting models).
Option C describes a database or storage service.
Option B describes a workflow agent or optimization AI.
(Reference: Google's training materials on Foundation Models define Diffusion Models as generative models that operate by gradually converting a state of random noise into a structured, meaningful output, most commonly for the generation of high-quality images and video.)

NEW QUESTION: 35
A data science team needs a centralized and organized location to store its various model versions, track their metadata, and easily deploy them to the respective applications. What Google Cloud service should they use?
A. Cloud Storage
B. Model Registry
C. BigQuery
D. Vertex AI Pipelines
Answer: (SHOW ANSWER)
A Model Registry (specifically part of Vertex AI Model Registry) is designed precisely for managing the lifecycle of machine learning models. It provides a centralized repository for storing, versioning, tracking metadata, and facilitating the deployment of models, which is essential for MLOps. Cloud Storage is for raw data, BigQuery for data warehousing, and Vertex AI Pipelines for workflow orchestration.
