

# ***KillTest***

高品質 より良いサービス



## 練習問題

<https://www.killtest.jp>

1年間無料アップデート

---

**Exam : Generative AI Leader**

**Title : Google Cloud Certified -  
Generative AI Leader Exam**

**Version : DEMO**

---

1. A company trains a generative AI model designed to classify customer feedback as positive, negative, or neutral. However, the training dataset disproportionately includes feedback from a specific demographic and uses outdated language norms that don't reflect current customer communication styles. When the model is deployed, it shows a strong bias in its sentiment analysis for new customer feedback, misclassifying reviews from underrepresented demographics and struggling to understand current slang or phrasing.

What type of model limitation is this?

- A. Data dependency
- B. Edge case
- C. Hallucination
- D. Overfitting

**Answer: A**

**Explanation:**

The core reason for the model's failure is that the training data itself was flawed (disproportionate demographic representation and outdated language). This flaw directly leads to the observed bias and poor performance on underrepresented groups and modern communication styles.

This is a classic example of Data Dependency, a fundamental limitation of all machine learning models, including generative AI. Data dependency refers to the absolute reliance of an AI model on the quality, completeness, and fairness of the data on which it was trained. Since the model essentially only mimics the patterns it learned from its dataset, if the dataset contains societal, demographic, or linguistic biases, the model will faithfully reproduce and amplify those biases in its output, leading to unfair classification for certain groups.

Hallucination (C) is the invention of facts or data.

Overfitting (D) is poor generalization because the model memorized the training data too well, typically resulting in very poor performance across all unseen data, not just specific demographics.

Bias is the result of the data dependency, not the fundamental limitation itself.

(Reference: Google's training on Generative AI Limitations identifies Data Dependency as the fundamental limitation where the model is limited by the scope and quality of its training data, directly leading to issues of bias when the data is not diverse or representative.)

2. 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: D**

**Explanation:**

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.)

3. 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: D**

**Explanation:**

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.

4. What does Vertex AI Search enable companies to do?

A. To index and retrieve information from the entire public web, providing a comprehensive view of publicly available data.

B. To surface the most popular and frequently accessed content based on global user search patterns and trends.

C. To compare products from numerous online retailers, allowing users to find the best deals and product options across the internet.

---

D. To ground LLM responses with first-party data, third-party data, and Google's knowledge graph.

**Answer: D**

**Explanation:**

Vertex AI Search is designed to enable powerful search experiences over an organization's own data (first-party), external data (third-party), and can leverage Google's knowledge graph to provide more relevant and accurate responses, especially when grounding Large Language Models (LLMs). It does not index the entire public web like Google Search.

5. A company is developing a generative AI application to analyze customer feedback collected through online surveys. Stakeholders are concerned about potential privacy risks associated with this data, as the feedback contains personally identifiable information (PII). They need to mitigate these risks before using the data to train the AI model.

What action should the company prioritize?

- A. Focusing on collecting only quantitative feedback data in future surveys.
- B. Ensuring that the AI model is trained on a large and diverse dataset.
- C. Implementing strong access controls to limit which teams can view the raw survey data.
- D. Applying data anonymization techniques to remove or obscure sensitive data.

**Answer: D**

**Explanation:**

The problem is the existence of Personally Identifiable Information (PII) within the customer feedback data, which introduces privacy risks for the development and training of the generative AI model. The goal is to mitigate these risks before using the data to train the AI model.

According to Google's Responsible AI and data handling best practices, when sensitive data like PII is present in a dataset intended for model training, the most critical step to prioritize is data minimization and privacy protection at the source. This is often achieved through anonymization or de-identification. Applying data anonymization techniques (D) directly addresses the risk by removing or obscuring the sensitive data elements. This prevents the PII from being embedded into the model's parameters during training, thereby eliminating the risk of data leakage or privacy violations in the AI application's outputs. This is a crucial early step in the ML lifecycle for datasets containing sensitive information.

Option C, implementing access controls, is a necessary security measure but is a reactive control that protects the raw data; it does not remove the PII risk from the derived model itself.

Option A is a long-term change to data collection but doesn't solve the problem for the existing data.

Option B relates to bias and accuracy, not specifically PII risk mitigation.

(Reference: Google Cloud's Secure AI Framework (SAIF) and Responsible AI principles emphasize protecting sensitive data at all stages of the ML lifecycle, with de-identification being the primary method before training.)