

Lab Summary: Using Gemini for Multimodal Retail Recommendations

Lab Info

Title: Getting Started with Vector Search and Embeddings

Learning Path: Advanced: Generative AI for Developers

Platform: Google Cloud Skills Boost

Duration: 1 hour

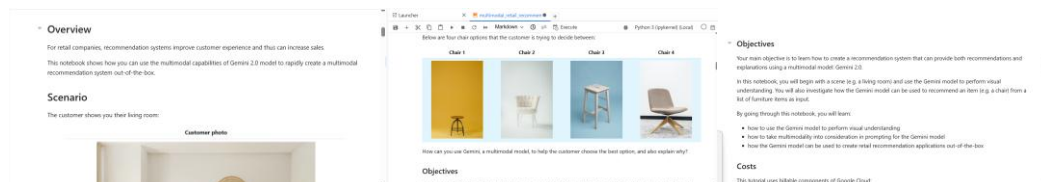
Objectives

The objective of this lab was to learn how to create a recommendation system using the multimodal capabilities of Gemini 2.0. This involved understanding how to input both images and text into the model to perform visual understanding and generate relevant recommendations for furniture in a retail scenario.

Lab Content

Task 1: Open the notebook in Vertex AI Workbench

I started by accessing the Vertex AI Workbench and opening the provided notebook that guides the lab activities. This environment supports running Python code and interacting with Google's AI models.



Task 2: Set up the notebook

In this step, I configured the necessary settings to run the Gemini model. This included importing libraries, setting up authentication for Google Cloud, and defining the image and text prompts that would be used in later steps.

Task 3: Use the Gemini model

I used the Gemini model to perform visual understanding of a customer's living room image. I combined the room image with a prompt asking the model to describe the room's visual elements and general atmosphere. The model responded with a detailed description, identifying walls, furniture, colors, and overall style.



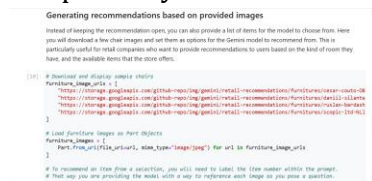
Task 4: Generate open recommendations based on built-in knowledge

Next, I asked Gemini to suggest new furniture for the room using only the image and the model's built-in knowledge. The model recommended a console table to be placed behind the sofa and explained how it fits the existing decor.



Task 5: Generate recommendations based on provided images

Finally, I provided a set of four chair images and asked Gemini to choose the most suitable one for the room. The model analyzed the room image and compared it to the chairs, explaining which chair matched best and why, based on visual harmony and style compatibility.



What I Learned

- I learned how to set up and use the Gemini 2.0 model within the Vertex AI Workbench.
- I gained experience in designing multimodal prompts that combine images and text.
- I understood how AI can describe visual scenes and make context-aware recommendations.
- I explored both open-ended recommendations and structured selection scenarios.
- I saw how AI can be used in retail to enhance customer experience through personalized product suggestions.