

Quick tip: Review the prerequisites before you run the lab

End Lab 00:52:37

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)

Open Google Cloud console

Username: student-04-95f7c37cafecl
Password: 7ffOrhm0AHSQ
Project ID: qwiklabs-gcp-01-18bfd71

Using Gemini for Multimodal Retail Recommendations

Lab 1 hour No cost Intermediate

★★★★★ Rate Lab

This lab may incorporate AI tools to support your learning.

Lab instructions and tasks

| | |
|---|--------|
| GSP1230 | 25/100 |
| Overview | |
| Objectives | |
| Setup and requirements | |
| Task 1. Open the notebook in Vertex AI Workbench | |
| Task 2. Set up the notebook | |
| Task 3. Use the Gemini model | |
| Task 4. Generate open recommendations based on built-in knowledge | |
| Task 5. Generate recommendations based on provided images | |
| Congratulations! | |

GSP1230



Overview

Gemini is a family of generative AI models developed by Google DeepMind that is designed for multimodal use cases.

For retail companies, recommendation systems improve customer experience and thus can increase sales. In this lab, you will learn how to use the Gemini model to rapidly create a multimodal recommendation system. The Gemini model can provide both recommendations and explanations using a multimodal model.

In this lab, you will begin with a scene (e.g. a living room) and use the Gemini model to perform visual understanding. You will also investigate how the Gemini model can be used to recommend an item (e.g. a chair) from a list of furniture items as input.

Gemini

Gemini is a family of powerful generative AI models developed by Google DeepMind, capable of understanding and generating various forms of content, including text, code, images, audio, and video.

Gemini API in Vertex AI

The Gemini API in Vertex AI provides a unified interface for interacting with Gemini models. This allows developers to easily integrate these powerful AI capabilities into their applications. For the most up-to-date details and specific features of the latest versions, please refer to the official [Gemini documentation](#).

Gemini Models

- [Gemini Pro](#): Designed for complex reasoning, including:

- Analyzing and summarizing large amounts of information.
- Sophisticated cross-modal reasoning (across text, code, images, etc.).
- Effective problem-solving with complex codebases.

- [Gemini Flash](#): Optimized for speed and efficiency, offering:

- Sub-second response times and high throughput.
- High quality at a lower cost for a wide range of tasks.
- Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompting for the Gemini model.
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompting for the Gemini model.
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompting for the Gemini model.
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompting for the Gemini model.
 - Effective problem-solving with complex codebases.

- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompts for the Gemini model
 - Sub-judgmental to super-modal reasoning across text, code, images, etc.
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompts for the Gemini model
 - Sub-judgmental to super-modal reasoning across text, code, images, etc.
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompts for the Gemini model
 - Sub-judgmental to super-modal reasoning across text, code, images, etc.
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.

- Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompting for the Gemini model, etc.
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompting for the Gemini model
 - Effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, sophisticated cross-modal reasoning (across text, code, images, etc.), and effective problem-solving with complex codebases.
- Gemini Flash: Optimized for speed and efficiency, offering:
 - Sub-second response times and high throughput.
 - High quality at a lower cost for a wide range of tasks.
 - Enhanced multimodal capabilities, including improved spatial understanding, new output modalities (text, audio, images), and native tool use (Google Search, code execution, and third-party functions).

Prerequisites

Before starting this lab, you should be familiar with:

- Basic Python programming.
- General API concepts.
- Running Python code in a Jupyter notebook on [Vertex AI Workbench](#).

Objectives

In this lab, you will learn how to:

- Use the Gemini model (`gemini-2.0-flash`) to perform visual understanding
- Take multimodality into consideration in prompting for the Gemini model