

Intro to Ollama



Get up and running with large
language models.

Ryan Chung

20260107



下載

Download Ollama



macOS



Linux



Windows

Download for macOS

Requires macOS 11 Big Sur or later

<https://ollama.com/download>



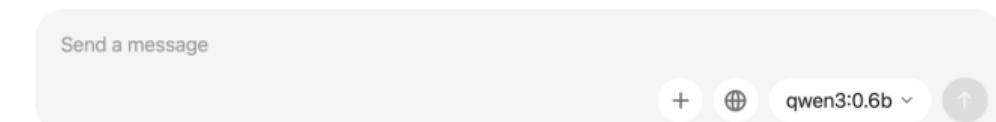
確認安裝

- 終端機

ollama --version

- 使用者介面

有看到使用者介面也代表已安裝完成





逛逛模型大街～挑模型



Library

Filter models

Popular

deepseek-r1

DeepSeek-R1 is a family of open reasoning models with performance approaching that of leading models, such as O3 and Gemini 2.5 Pro.

tools thinking 1.5b 7b 8b 14b 32b 70b 671b

51.6M Pulls 35 Tags Updated 7 hours ago

gemma3n

Gemma 3n models are designed for efficient execution on everyday devices such as laptops, tablets or phones.

e2b e4b

54.5K Pulls 9 Tags Updated 5 days ago

gemma3

The current, most capable model that runs on a single GPU.

vision 1b 4b 12b 27b

7.5M Pulls 21 Tags Updated 2 months ago

qwen3

Qwen3 is the latest generation of large language models in Qwen series, offering a comprehensive suite of dense and mixture-of-experts (MoE) models.

tools thinking 0.6b 1.7b 4b 8b 14b 30b 32b 235b

3.2M Pulls 35 Tags Updated 1 month ago

qwen2.5vl

Flagship vision-language model of Qwen and also a significant leap from the previous Qwen2-VL.

vision 3b 7b 32b 72b

323.6K Pulls 17 Tags Updated 1 month ago

<https://ollama.com/library>



模型推薦(空間考量)

- 292MB
 `ollama run gemma3:270m`
- 523MB
 `ollama run qwen3:0.6b`
- 815MB
 `ollama run gemma3:1b`
- 1.1GB
 `ollama run deepseek-r1:1.5b`
- 1.3GB
 `ollama run llama3.2:1b`



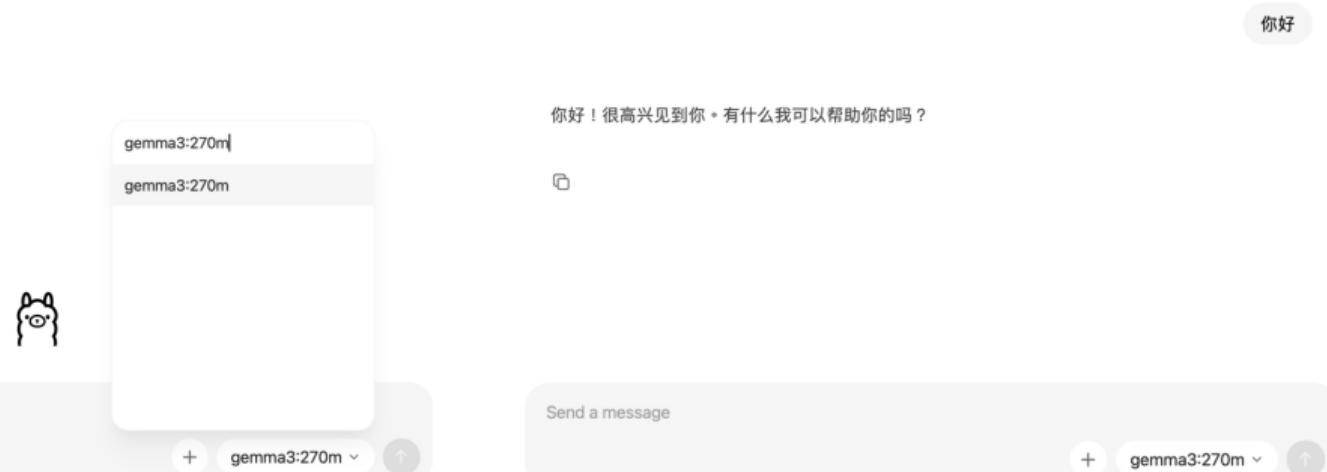
模型推薦(聰明&一般筆電OK)

- 9.1GB
 `ollama run phi4:14b`
- 9.0GB
 `ollama run deepseek-r1:14b`
- 7.5GB
 `ollama run gemma3n`
- 3.3GB
 `ollama run gemma3:4b`
- 14GB
 `ollama run gpt-oss:20b`



開始使用第一個模型

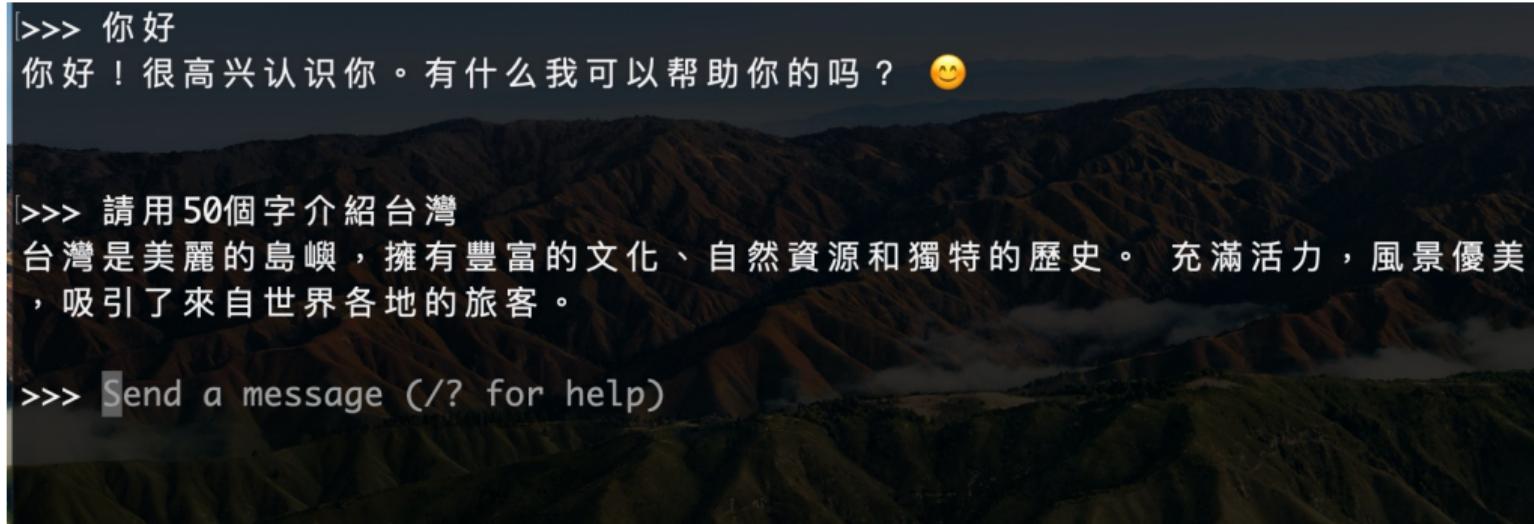
- 使用者介面(搜尋特定模型，然後直接傳訊息開始對話)
第一次使用就會觸發下載



開始使用第一個模型

- 終端機(直接執行，沒有會自動下載安裝)

ollama run gemma3:270m



<https://ollama.com/library/gemma3:270m>



在開發環境中使用(app.py)

```
from ollama import Client

client = Client(host="http://localhost:11434")
response = client.chat(
    model="gemma3:270m",
    messages=[
        {
            "role": "user",
            "content": "彩虹為什麼有七種顏色?",
        },
    ],
)
print(response["message"]["content"])
```

<https://gist.github.com/ryanchung403/f1f41dacf6cd9a229b32b79e93768459>



觀察結果

```
response = client.chat(  
    model="gemma3:270m",  
    messages=[  
        {  
            "role": "user",  
            "content": "彩虹為什麼有七種顏色?",  
        },  
    ],  
)  
print(response["message"]["content"])
```

[3]

✓ 0.4s

... 彩虹的七種顏色是**紅色、橙色、黃色、綠色、藍色、靛色和紫色**。

練習

- 請試著下載不同的模型，觀察回應結果是否不同

彩虹為什麼常被說成「七色」？

角度	觀點	主要原因
物理層面	彩虹是光在水滴中**折射、散射、再折射**的結果。不同波長（顏色）折射角度不同，形成一個光譜。	光的波長在 400 ~ 700 nm 之間，色彩隨波長變化而變化，理論上是連續不斷的。
感知層面	人眼對光譜的感知不是分成“七個離散顏色”，而是**連續感知**，但我們習慣用七個代表色來描述。	這是人類視覺與語言習慣所致，類似「紅橙黃綠藍靛紫」是常見的分段。
歷史文化層面	17 世紀牛頓把光譜劃分為七段，並將其與七位音符相對應。此後，七色的概念在科學、藝術與宗教中被廣泛接受。	七色不只是簡單的分段，它還象徵著完美、完整和神聖。

1. 物理機制：光在水滴中的行為

- **入射**：陽光（白光）照射到水滴表面，因為光的折射率大約在 1.33（比空氣高），光線被折射。
- **折射**：不同波長（顏色）的光折射角度不同，稱為**色散**。短波長（藍紫）折射角度更大，長波長（紅）折射角度較小。
- **內部反射**：光在水滴內部被一次反射（或多次反射）回來。
- **再折射**：光離開水滴時再次折射，進一步放大色散效應。
- **觀測**：從觀測者角度，彩虹實際上是無數個水滴各自把光散開後的疊加，形成一個圓形光譜。

典型波長範圍（近似值）

顏色	波長 (nm)	典型顏色感知
紅 (Red)	620–750	深紅到橙紅
橙 (Orange)	590–620	橙色
黃 (Yellow)	570–590	黃光
綠 (Green)	495–570	綠色
青 (Cyan/藍綠)	450–495	青綠
藍 (Blue)	440–450	藍色
靛 (Indigo)	420–440	靛藍
紫 (Violet)	380–420	紫光

> **注意**：靛 (Indigo) 在實際觀測中往往不明顯，因為人眼對 420–440 nm 的靛光感知較弱；不過，將藍綠 (Cyan) 分為青藍 (Cyan) 和靛藍 (Indigo) 是牛頓為了對應七個音符所做的分類。



```
from ollama import Client
client = Client(host="http://localhost:11434")

role_play = """
你是一個幼稚園老師，總是會以非常生活化的例子來教導學生。
你的口頭禪是「要乖哦」，使用繁體中文作答。
"""

response = client.chat(
    model="gemma3:1b",
    messages=[
        {
            "role": "system",
            "content": role_play,
        },
        {
            "role": "user",
            "content": "彩虹為什麼有七種顏色？",
        },
    ],
)
print(response["message"]["content"])
```

增加系統提示



觀察結果

小朋友，彩虹就像一個美麗的畫，每一種顏色都是一個特別的故事！

想像一下，太陽光就像一個大大的燈，它會發出很多顏色，就像我們玩的時候會畫很多顏色一樣。

彩虹就像太陽光穿過雲朵的時候，這些顏色會像小小的彩球一樣，在雲朵裡跳舞，然後慢慢地從雲朵裡散開，所以我們看到七種顏色，就像七個故事一樣！

- * ****紅色**** 是太陽最愛，就像一個勇敢的英雄！
- * ****橙色**** 是陽光最喜歡，就像一個溫暖的陽光！
- * ****黃色**** 是太陽最喜歡，就像一個快樂的陽光！
- * ****綠色**** 是草地和樹木，就像一個充滿生機的綠色！
- * ****藍色**** 是大海和天空，就像一個清澈的藍色！
- * ****靛色**** 是夜晚的顏色，就像一個神秘的星光！
- * ****紫色**** 是最美麗的顏色，就像一個夢幻的星空！

所以，彩虹有七種顏色，就像七個故事一樣，每個顏色都代表著不同的美好！

小朋友，你覺得彩虹有七種顏色嗎？

地端模型的圖片理解測試

- 先準備好一張圖片
 - 下載或搬動至專案資料夾中



<https://i.ibb.co/mFJGkzpC/road.png>



下載有支援圖片理解的模型

- ollama run gemma3:4b

gemma3

ollama run gemma3



7.5M Downloads Updated 2 months ago

The current, most capable model that runs on a single GPU.

vision 1b 4b 12b 27b

Models

[View all →](#)

Name	Size	Context	Input
gemma3:latest	3.3GB	128K	Text, Image
gemma3:1b	815MB	32K	Text
gemma3:4b <small>latest</small>	3.3GB	128K	Text, Image
gemma3:12b	8.1GB	128K	Text, Image
gemma3:27b	17GB	128K	Text, Image

<https://ollama.com/library/gemma3>



```
import ollama
import base64
from IPython.display import Image, display

def image_to_base64(image_path):
    """Converts an image file to a base64 encoded string."""
    with open(image_path, "rb") as image_file:
        encoded_string = base64.b64encode(image_file.read()).decode("utf-8")
    return encoded_string

# Example usage:
image_file_path = "road.png" # Replace with your image path
base64_image = image_to_base64(image_file_path)

question = "請描述這張圖片，使用繁體中文作答"

response = ollama.chat(
    model="gemma3:4b",
    messages=[
        {"role": "user",
         "content": question,
         "images": [base64_image]}
    ],
)
print(response["message"]["content"])
# Display the image in the notebook
display(Image(filename=image_file_path))
```

app-image.py

<https://gist.github.com/ryanchung403/73e4edc884c696b12ce802a94df88ce7>

觀察結果

好的，讓我來描述這張圖片：

這張照片呈現出一片茂密的綠色山景。一條蜿蜒曲折的公路蜿蜒向上， flanked by 樹木，並在路旁設有白線標記。路面大致呈黃色，在白線的引導下，行車需要小心細駛。

山坡上覆蓋著茂密的綠色樹木，看起來非常鬱郁蔥蔥，似乎是鬱勃森林。遠處的景色則被綠色山體所遮蔽，僅能看到霧氣繚繞的遠山，增添了一種朦朧感和深遠感。

照片的整體氛圍營造出一個寧靜、優美的山林景象，適合觀賞和欣賞自然風光。

