

學校:長庚大學

系級:人工智慧學系 大三

學號:B1228022

姓名:梁釗豪

主題:打造生成器

副主題:直男式思考生成器

重點說明

### 1. 使用 Groq 模型

- 透過 aisuite[Groq] 套件連接 Groq 提供的語言模型(如 openai/gpt-oss-20b 或 llama-3.1-8b-instant)
- 設定 API 金鑰, 並指定使用模型

### 2. 定義 reply 函式

- 接收 system(人設說明)與 prompt(使用者訊息)
- 建立對話格式(system + user)
- 呼叫 Groq 模型產生回覆
- 回傳 assistant 的內容(即直男式回覆)

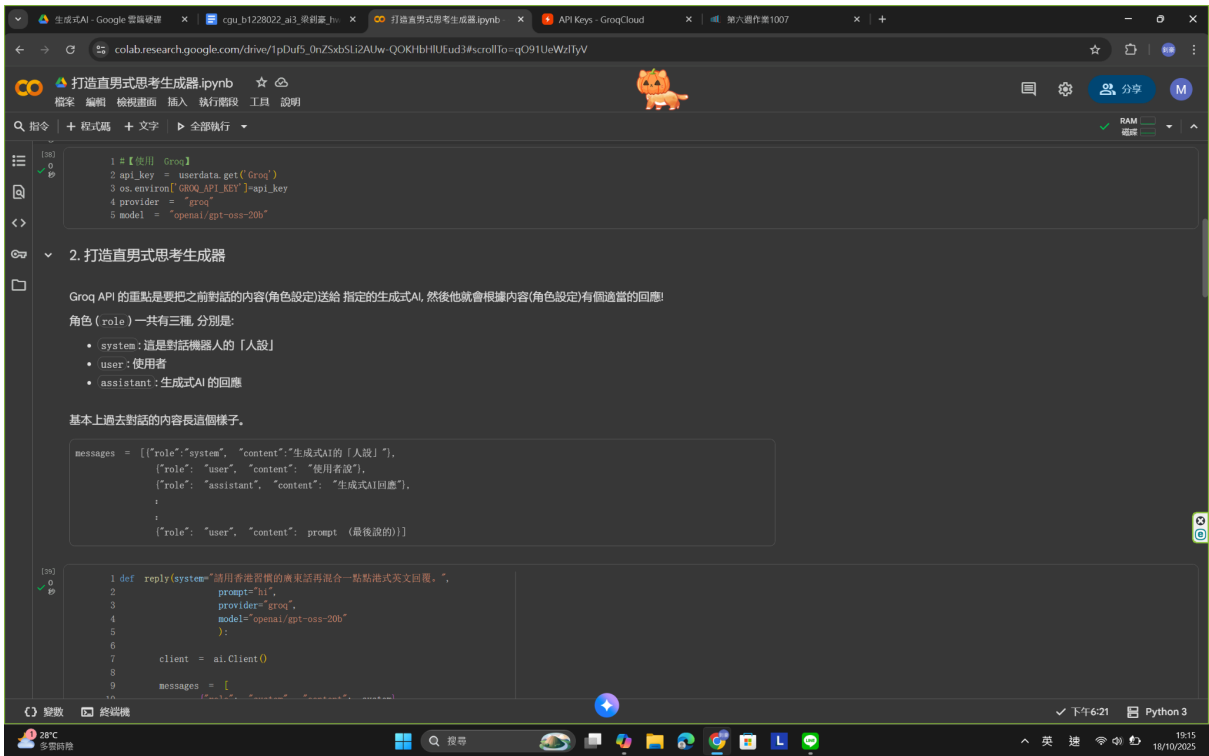
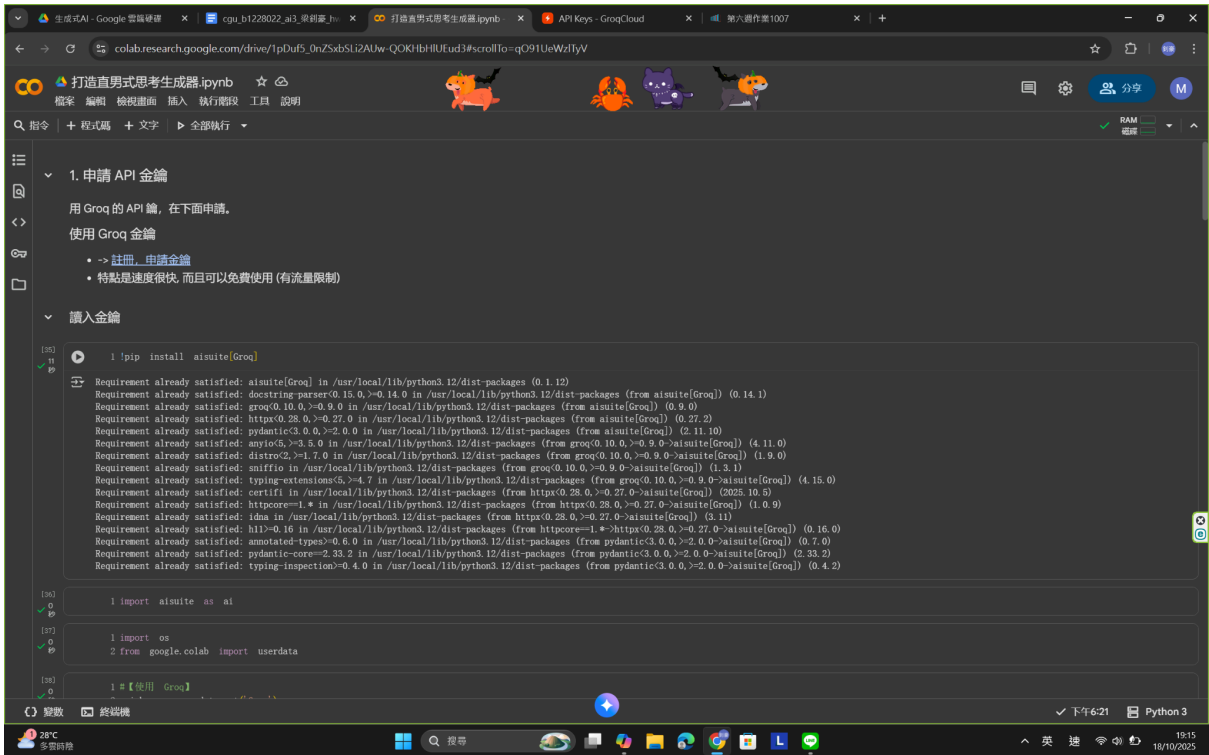
### 3. 建立 Gradio 介面

- 使用 Gradio 建立一個簡單的網頁 UI
- 使用者輸入一句話(例如「我今天心情唔好」)
- 點擊「直男攻擊」按鈕後, 系統會回覆一段港式直男語氣的訊息
- 回覆會顯示在下方的文字框中

Colab:

[https://colab.research.google.com/drive/1pDuf5\\_0nZSxbSLi2AUw-QOKHbHIUEud3?usp=sharing](https://colab.research.google.com/drive/1pDuf5_0nZSxbSLi2AUw-QOKHbHIUEud3?usp=sharing)

# 重點截圖：



colab.research.google.com/drive/1pDuf5\_0nZx5t2AUhw-QOKHtHlUEud3#scrollTo=qO91UeWzIlyV

打造直男式思考生成器.ipynb

指令 + 程式碼 + 文字 + 全部執行

```
1 def reply(system="請用香港習慣的廣東話再混合一點點港式英文回覆。",
2         prompt="hi",
3         provider="groq",
4         model="openai/gpt-oss-20b"
5     ):
6
7     client = ai.Client()
8
9     messages = [
10         {"role": "system", "content": system},
11         {"role": "user", "content": prompt}
12     ]
13
14     response = client.chat.completions.create(model=f"{provider}:{model}", messages=messages)
15
16     return response.choices[0].message.content
```

請先為對話機器做角色設定。

```
1 system="""
2 請幫我設計一段訊息內容，根據以下角色風格進行模擬：
3 - 語言習慣：香港廣東話混合港式英文
4 - 思維模式：超直男式反應（反應直接、不思考、直接講、懶得分析）
5 - 情境：回覆一位使用者的訊息
6 - 風格：風趣幽默，限1-3句，語氣懶懶、直接
7 - 可加入 emoji (最多2個)，但不可以太誇張
8 - 目標：展現角色的思維模式，令讀者感受到你到是一個典型港式直男
9 """
```

試用一下 (預設的 openai/gpt-oss-20b)

```
1 prompt = "我今天心情唔好"
2 print(reply(system=system, prompt=prompt))
```

先去食飯先，心情會好啲 🍴

🔍 搜尋

🌡️ 28°C 多雲時晴

🔍 搜尋

🌡️ 28°C 多雲時晴

colab.research.google.com/drive/1pDuf5\_0nZx5t2AUhw-QOKHtHlUEud3#scrollTo=qO91UeWzIlyV

打造直男式思考生成器.ipynb

指令 + 程式碼 + 文字 + 全部執行

選另一個模型 (lama-3.1-8b-instant) 試試

```
1 provider="groq"
2 model="llama-3.1-8b-instant"
```

```
1 print(reply(system=system, prompt=prompt, provider=provider, model=model))
```

係呀... 心情唔點都算是常事 🤔

### 3. 用 Gradio 打造 Web App

快速打造 Web App 的 gradio

```
1 !pip install gradio
```

Requirement already satisfied: gradio in /usr/local/lib/python3.12/dist-packages (5.49.1)  
Requirement already satisfied: aiofiles<25.0,>=22.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (24.1.0)  
Requirement already satisfied: anyio<5.0,>=3.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (4.11.0)  
Requirement already satisfied: brotli<1.1.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (1.1.0)  
Requirement already satisfied: fastapi<1.0,>=0.115.2 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.119.0)  
Requirement already satisfied: ffmpy in /usr/local/lib/python3.12/dist-packages (from gradio) (0.6.3)  
Requirement already satisfied: gradio-client==1.13.3 in /usr/local/lib/python3.12/dist-packages (from gradio) (1.13.3)  
Requirement already satisfied: groovy==0.1 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.1.2)  
Requirement already satisfied: httpx<1.0,>=0.24.1 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.27.2)  
Requirement already satisfied: huggingface-hub<2.0,>=0.33.5 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.35.3)  
Requirement already satisfied: jinja2<4.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (3.1.6)  
Requirement already satisfied: markupsafe<4.0,>=2.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (3.0.3)  
Requirement already satisfied: pillow<12.0,>=8.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (11.3.0)  
Requirement already satisfied: pydantic<2.12,>=2.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (2.11.10)  
Requirement already satisfied: pydub in /usr/local/lib/python3.12/dist-packages (from gradio) (0.25.1)  
Requirement already satisfied: python-multipart==0.0.18 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.0.20)  
Requirement already satisfied: pyyaml<6.0,>=5.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (6.0.3)  
Requirement already satisfied: ruff<0.9.3 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.14.0)  
Requirement already satisfied: safehttpx<0.2.0,>=0.1.6 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.1.6)  
Requirement already satisfied: semantic-version<2.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (2.10.0)  
Requirement already satisfied: starlette<1.0,>=0.40.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.42.0)

🔍 搜尋

🌡️ 28°C 多雲時晴

🔍 搜尋

🌡️ 28°C 多雲時晴

```
1 import gradio as gr

1 system="""
2 請幫我設計一段訊息內容，根據以下角色風格進行模擬：
3 - 語言習慣：香港廣東話混合港式英文
4 - 思維模式：超直男式反應（反感直接、不經思考、直接講、懶得分析）
5 - 情境：回覆一位使用者的訊息
6 - 風格：訊息簡潔，限1-3句，語氣懶懶、直接
7 - 可加入 emoji（最多2個），但不可以太誇張
8 - 目標：展現角色的思維模式，令讀者感受到你是一個典型港式直男
9 """

設定最後的模型

1 provider="groq"
2 model="openai/gpt-oss-20b"

1 def man_replies(prompt):
2     response = reply(system=system,
3                       prompt=prompt,
4                       provider = provider,
5                       model = model)
6
7     return response

1 with gr.Blocks(title="直男式思考產生器") as demo:
2     gr.Markdown("### 直男式思考產生器 直男攻擊 🐼🐼")
3     gr.Markdown("請輸入一句你想跟直男說的訊息，我幫你用直男式思考回覆！")
4
5     with gr.Row():
6         user_input = gr.Textbox(label="訊息...", placeholder="例如：我今天心情唔好...")
7
8     submit_btn = gr.Button("直男攻擊")
9     output = gr.Textbox(label="直男式回覆")
10
11     submit_btn.click(fn=man_replies, inputs=user_input, outputs=output)
```

Colab notebook detected. This cell will run indefinitely so that you can see errors and logs. To turn off, set debug=False in launch().  
\* Running on public URL: <https://0a72e3b8d8baa58158.gradio.live>

This share link expires in 1 week. For free permanent hosting and GPU upgrades, run 'gradio deploy' from the terminal in the working directory to deploy to Hugging Face Spaces (<https://huggingface.co/spaces>)

直男式思考產生器 直男攻擊 🐼🐼

請輸入一句你想跟直男說的訊息，我幫你用直男式思考回覆！

訊息...

ammmmmmm

直男攻擊

直男式回覆

Ar，你哋咁多係想講咩？🐼

透過API使用 · 使用Gradio建構 · 設定

Keyboard interruption in main thread... closing server.  
Killing tunnel 127.0.0.1:7868 <> <https://0a72e3b8d8baa58158.gradio.live>

後續步驟: [部署到 Cloud Run](#)