

GPT4all ReadMe

Anmol Valecha

June 2024

Introduction

This document provides a step-by-step guide for downloading, installing, and running the GPT4all model on an M1 Mac using Flask for the web interface.

Prerequisites

- M1 Mac
- Python installed
- Visual Studio Code (VS Code) installed
- Internet connection

Step-by-Step Instructions

1. Setting Up Python Virtual Environment

1. Create a New Folder:

- Create a new folder named `GPT4all_Assignment` and navigate into it:

```
mkdir GPT4all_Assignment  
cd GPT4all_Assignment
```

2. Create a Virtual Environment:

- Run the following command to create a virtual environment:

```
python3 -m venv .venv
```

3. Activate the Virtual Environment:

- Activate the virtual environment with:

```
source .venv/bin/activate
```

4. Install GPT4all:

- Install GPT4all using pip in the activated virtual environment:

```
pip install gpt4all
```

5. Install Flask:

- Install Flask using pip in the activated virtual environment:

```
pip install flask
```

2. Download User Interface

- Download the GPT4all user interface from the official website: <https://gpt4all.io/>
- Follow the installation instructions provided on the website to set up the user interface on your local machine.

3. Setting Up Flask Web Interface

1. Create Flask Application:

- In your project directory, create a Python file (e.g., 'app.py') for your Flask application.

2. Write Flask Application Code:

- Implement Flask routes and logic to interact with the GPT4all model. Here's an example based on your provided code:

Listing 1: app.py

```
from flask import Flask, request, jsonify
from gpt4all import GPT4All

app = Flask(__name__)
model = GPT4All(model_name='orca-mini-3b-gguf2-q4_0.gguf')

@app.route('/api/generate', methods=['POST'])
def generate():
    try:
        data = request.json
        prompt = data.get('prompt')
        if not prompt:
            return jsonify({"error": "No-prompt-provided"}), 400
```

```

        with model.chat_session():
            response = model.generate(prompt, temp=0)

        return jsonify({"response": response})

    except Exception as e:
        return jsonify({"error": str(e)}), 500

@app.route('/', methods=['GET'])
def hello():
    return "Hello, World!"

if __name__ == '__main__':
    app.run(debug=True, port=5050)

```

3. Run Flask Application:

- Start the Flask server:

```
python app.py
```

4. Access the Web Interface:

- Open a web browser and go to 'http://localhost:5050' to interact with the GPT4all model through the Flask interface.

4. Interacting with GPT4all

• Using Curl:

- Use the following curl command on another terminal to interact with GPT4all via Flask:

```

curl -X POST http://localhost:5050/api/generate \
-H "Content-Type: application/json" \
-d '{"prompt": "Your Question"}'

```

• Using User Interface:

- Open the installed GPT4all user interface on your machine.
- Select the desired model (e.g., gpt4all, phi3) from the interface.
- Click on "New Chat" to start a new chat session and interact with the model.

Example Chat Interactions

```
Assignment3 -- zsh -- 115x38

(gpt4all-venv) annolvaalecha@Annolis-MacBook-Air Assignment3 % curl -X POST http://localhost:5050/api/generate \
-H "Content-Type: application/json" \
-d '{"prompt": "Hello"}'

{
  "response": " Hello! How may I assist you today?"
}
(gpt4all-venv) annolvaalecha@Annolis-MacBook-Air Assignment3 % curl -X POST http://localhost:5050/api/generate \
-H "Content-Type: application/json" \
-d '{"prompt": "Name a few countries in Asia"}'

{
  "response": " Sure, here are some countries in Asia:\n1. China\n2. India\n3. Japan\n4. South Korea\n5. Thailand\n6. Vietnam\n7. Indonesia\n8. Malaysia\n9. Philippines\n10. Singapore"
}
(gpt4all-venv) annolvaalecha@Annolis-MacBook-Air Assignment3 % curl -X POST http://localhost:5050/api/generate \
-H "Content-Type: application/json" \
-d '{"prompt": "Tell me something more about India"}'

{
  "response": " India is a country located in South Asia and is known for its diverse culture, rich history, and vibrant economy. It is the seventh-largest country in the world by both area and population. The official language of India is Hindi, but there are many other languages spoken throughout the country, including English, Bengali, Tamil, Telugu, Marathi, and more. \n\nIndia is known for its delicious cuisine, which includes dishes such as butter chicken, biryani, and naan. It is also famous for its traditional textiles, jewelry, and art. India has a rich history that dates back thousands of years, with many ancient ruins and monuments still visible today. Some of the most well-known landmarks in India include the Taj Mahal, Agra Fort, and the Red Fort in Delhi. \n\nIndia is also known for its diverse culture, which includes traditions such as yoga, meditation, and spiritual practices like Hinduism, Buddhism, and Sikhism"
}
(gpt4all-venv) annolvaalecha@Annolis-MacBook-Air Assignment3 %
```

Figure 1: Example Chat Interaction with GPT4all via Flask

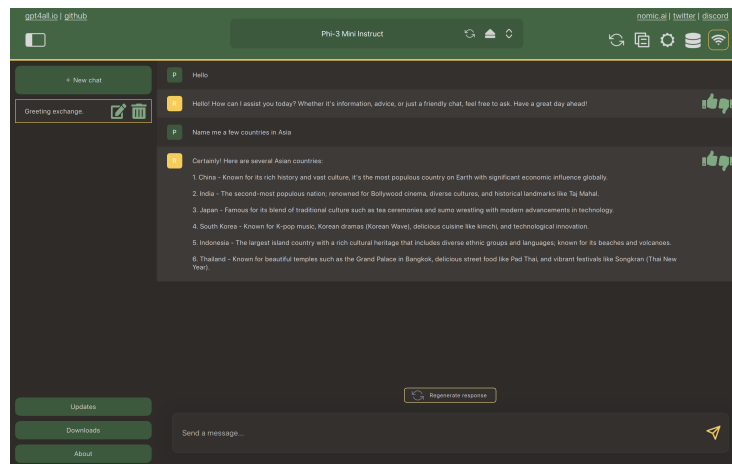


Figure 2: Example Chat Interaction with GPT4all via UI