There are two popular tools that support running LLMs on your PC:

**1. LM Studio** (watch tutorial at <https://www.youtube.com/watch?v=ygUEbCpOOLg> )

Download it at <https://lmstudio.ai/> and then follow the video tutorial to download and load LLMs.

Easy, but does not support calling cloud APIs from OpenAI and other companies.

**2. Ollama** (watch tutorial at <https://www.youtube.com/watch?v=BzFafshQkWw&t=496s> )

a. Download Ollama at <https://ollama.com/download> or <https://github.com/ollama/ollama>

b. Go to <https://ollama.com/search> and install models you’d like to try.

For instance, to install and run the smallest DeepSeek R1 model, you can type “ollama run deepseek-r1:1.5b” in command line (Powershell for Windows or Terminal for Mac). Once the download is completed, you can chat with the model

A screenshot of a computer

Description automatically generated

To exit the chat, press CTRL + D.

c. (Optional) Install Open WebUI to get a browser interface

i) Install Docker Desktop from <https://www.docker.com/products/docker-desktop/>

Open the Docker Desktop app after installation. If you received an error related to “WSL”, try typing “wsl --update” in your command line to update it manually.

ii) Install Open WebUI (<https://github.com/open-webui/open-webui>) with the following command

docker run -d -p 3000:8080 --add-host=host.docker.internal:host-gateway -v open-webui:/app/backend/data --name open-webui --restart always ghcr.io/open-webui/open-webui:main

After installation, you can access Open WebUI at <http://localhost:3000>.

d. (Optional) Using OpenAI API (tutorial at <https://www.youtube.com/watch?v=ZUc50fcWO2E>)

Get the API key from <https://openai.com/api/> (needs to deposit $5 into your account)