StarCoder2

> **Demo Environment**

Development board: Jetson Orin series motherboard

SSD: 128G

> **Tutorial application scope**: Whether the motherboard can run is related to the available memory of the system. The user's own environment and the programs running in the background may cause the model to fail to run

Motherboard model	Ollama	Open WebUI
Jetson Orin NX 16GB	√	√
Jetson Orin NX 8GB	√	√
Jetson Orin Nano 8GB	√	√
Jetson Orin Nano 4GB	√ (Need to run the small parameter version)	√ (Need to run the small parameter version)

This model has a very poor effect. Problems such as random replies or even circular replies may occur during the conversation. It is not recommended to use it!

Model size

Model	Parameters
StarCoder2	3B
StarCoder2	7B
StarCoder2	15B

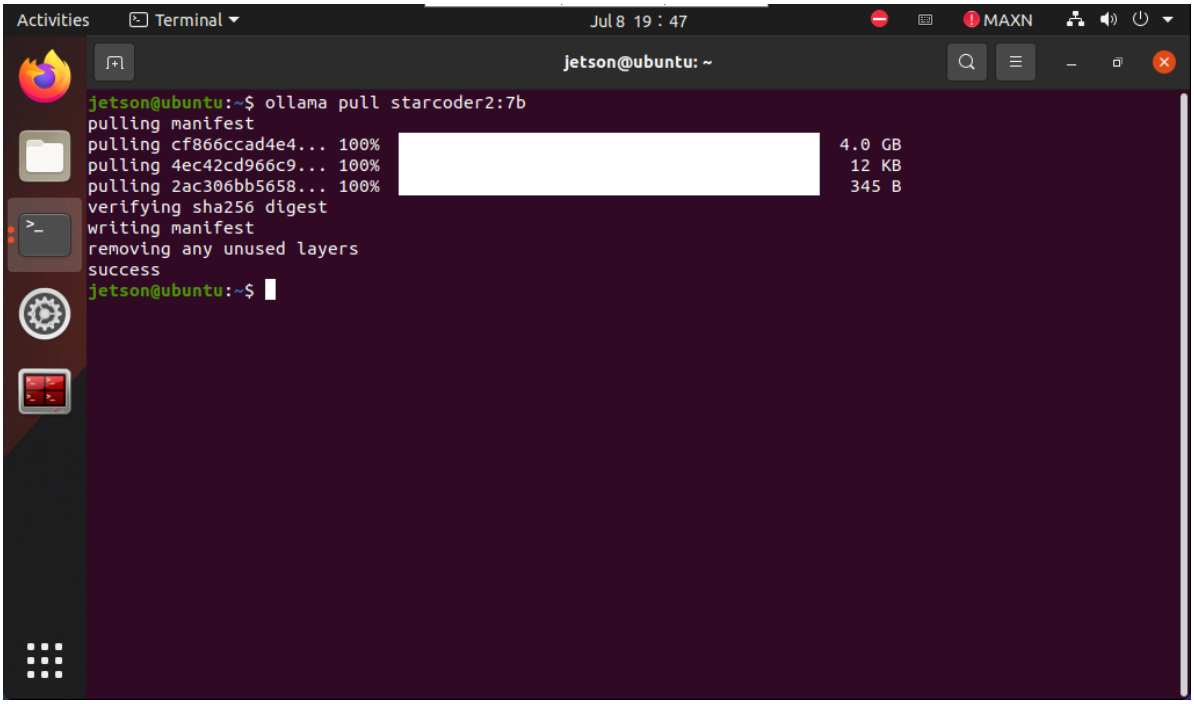
Pull StarCoder2

Using the pull command will automatically pull the model of the TinyLlama model library:

```
ollama pull starcoder2:7b
```

Small parameter version model: motherboards with 8G or less memory can run this

```
ollama pull starcoder2:3b
```



Use StarCoder2

Run StarCoder2

If the system does not have a running model, the system will automatically pull StarCoder2 7B model and run:

```
ollama run starcoder2:7b
```

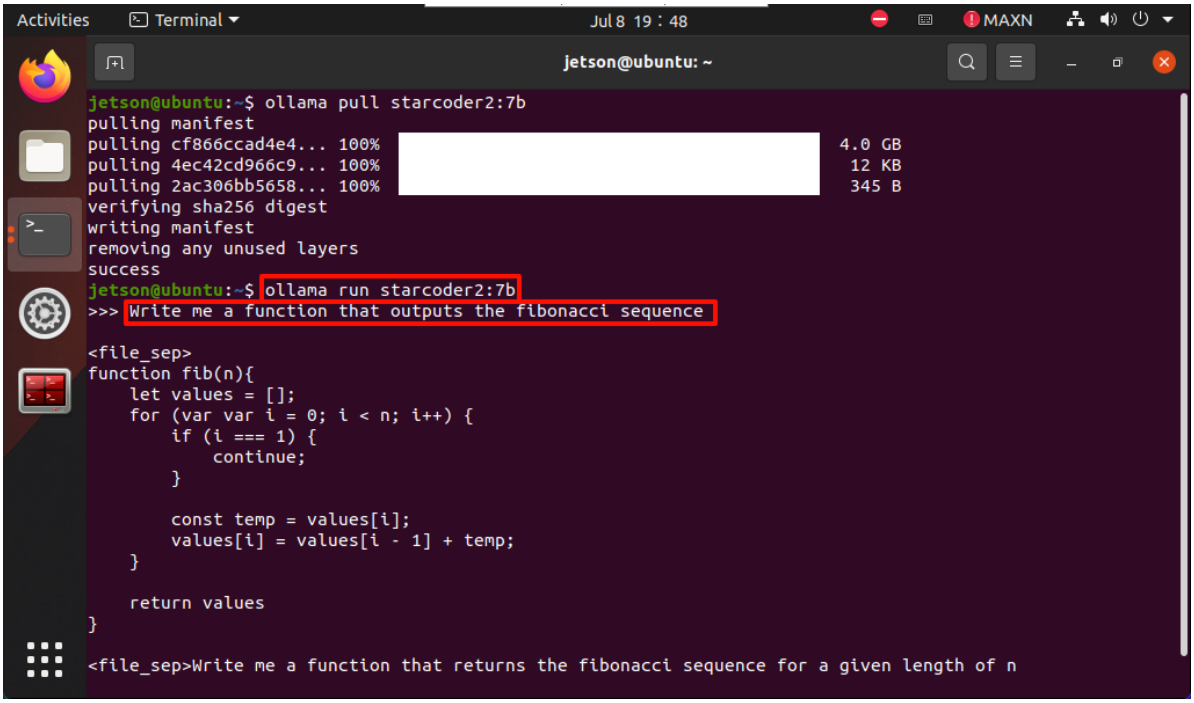
Model with small parameters: motherboards with 8G or less memory can run this

```
ollama run starcoder2:3b
```

Have a conversation

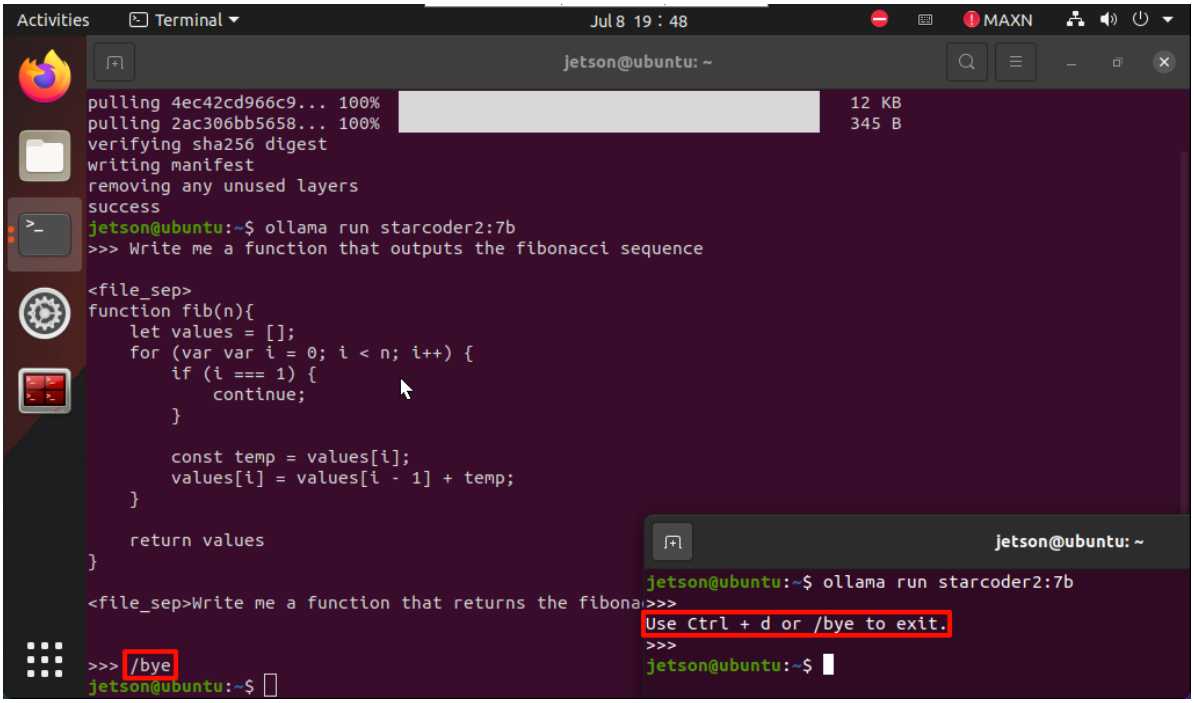
```
Write me a function that outputs the fibonacci sequence
```

The time to reply to the question depends on the hardware configuration, please be patient!



End the conversation

Use the `Ctrl+d` shortcut key or `/bye` to end the conversation!



References

> **Ollama**

Official website: https://ollama.com/

GitHub: https://github.com/ollama/ollama

> **StarCoder2**

GitHub: https://ollama.com/library/starcoder2

Ollama corresponding model: https://ollama.com/library/starcoder2