Gemma model

Gemma model

Model scale

Got Gemma

Run Gemma

Run Gemma

Dialogue

End conversation

Reference material

Demonstration environment

Development board: Raspberry Pi 5B (8G RAM)

SD(TF)card: 64G (Above 16G, the larger the capacity, the more models can be experienced)

```
Raspberry Pi 5B (8G RAM): Run 8B and below parameter models
Raspberry Pi 5B (4G RAM): Run 3B and below parameter models
Raspberry Pi 5B (2G RAM): Run 3B and below parameter models, can't run Gemma
model
```

Gemma is a new open model developed by Google and its DeepMind team.

Model scale

Model	Parameter
Gemma	2B
Gemma	7B

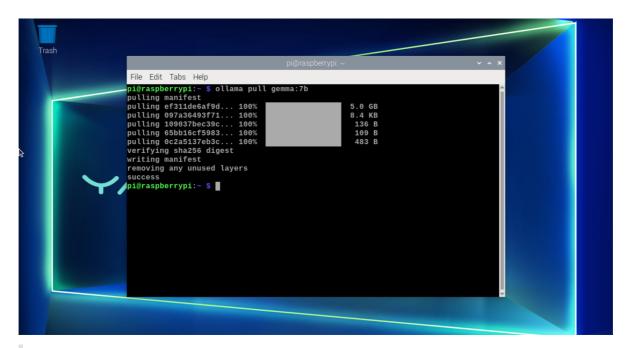
```
Raspberry Pi 5B (8G RAM): Test Gemma model with 7B parameters!
Raspberry Pi 5B (4G RAM): Test Gemma Model with 2B Parameters!
```

Got Gemma

Using the pull command will automatically retrieve the models from the Ollama model library.

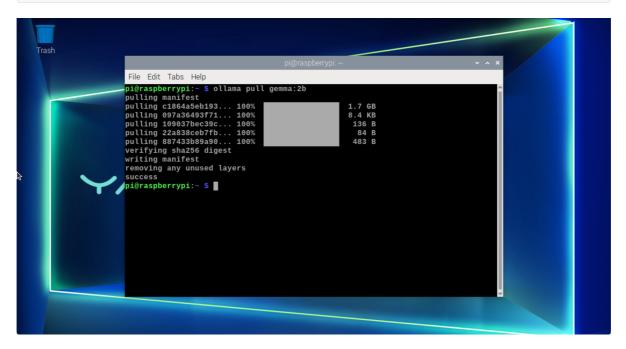
Raspberry Pi 5B (8G RAM)

```
ollama pull gemma:7b
```



Raspberry Pi 5B (4G RAM)

ollama pull gemma:2b



Run Gemma

Run Gemma

Raspberry Pi 5B (8G RAM)

If the system does not have a running model, the system will automatically pull the Gemma 7B model and run it: **The factory image does not provide this image, you need to download it yourself!**

ollama run gemma:7b

Raspberry Pi 5B (4G RAM)

If the system does not have a running model, the system will automatically obtain the Gemma 2B model and run it.

```
ollama run gemma:2b
```

Dialogue

Raspberry Pi 5B (8G RAM)

```
Introduce yourself in Chinese

print Helloworld in C

print Helloworld in python
```

The time to reply to the question is related to the hardware configuration, please be patient.

```
pi@raspberrypi:~ $ |ollama run gemma:7b|
>>> Introduce yourself in Chinese|
你好!我叫模型,一个语言模型,可以帮助你完成各种文字相关的任务。请何有什么我能帮忙的吗?
>>> print HelloWorld in C
c
#include <stdio.h>
int main() {
    printf("HelloWorld!\n");
    return 0;
}
>>> print HelloWorld in python
    python
    printf("HelloWorld!")
>>> Send a message (/? for help)
```

Raspberry Pi 5B (4G RAM)

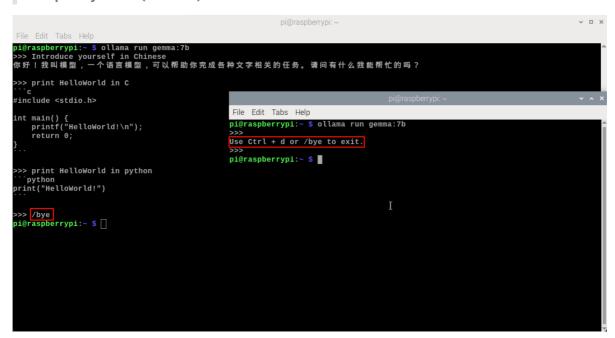
```
print HelloWorld in C
```

The time to reply to the question is related to the hardware configuration, please be patient.

End conversation

You can end the conversation by using the shortcut key 'Ctrl+d' or '/bye'.

Raspberry Pi 5B (8G RAM)



Raspberry Pi 5B (4G RAM)

Reference material

Ollama

Website: https://ollama.com/

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

Gemma

GitHub: https://github.com/google-deepmind/gemma

Ollama model: https://ollama.com/library/gemma