WizardLM-2

WizardLM-2

- 1. Model scale
- 2. Pull WizardLM-2
- 3. Use WizardLM-2
 - 3.1. Run WizardLM-2
 - 3.2. Have a conversation
 - 3.3. End the conversation

References

Demonstration environment

Development board: rdk x5 motherboard

sd card: 64G

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

Board model	Ollama
rdk x5 8GB	√
rdk x5 4GB	×

WizardLM-2 is Microsoft's latest advanced large-scale language model, which improves performance in complex dialogue, multi-language, reasoning, and intelligent agent capabilities.

1. Model scale

Model	Parameter
WizardLM-2	7B

2. Pull WizardLM-2

Pull the following parameter models for the RDK X5 8Gb motherboard

Using the pull command will automatically pull the model from the Ollama model library:

ollama pull wizardlm2:7b

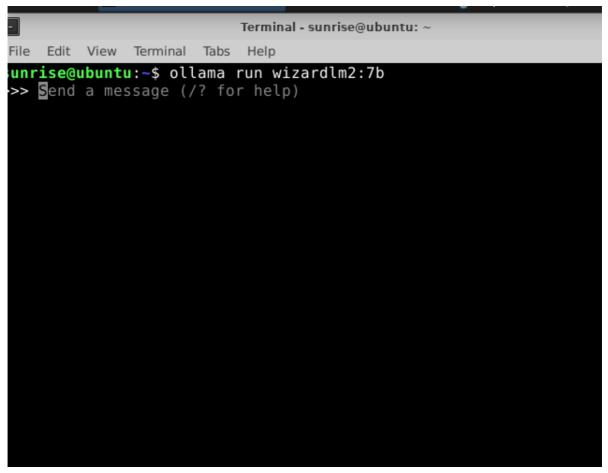
```
Terminal - sunrise@ubuntu: ~
File Edit View Terminal Tabs Help
sunrise@ubuntu:~$ ollama pull wizardlm2:7b
pulling manifest
pulling 431800fe7a30... 100%
                                               4.1 GB
pulling 43070e2d4e53... 100%
                                                11 KB
pulling 245f91e42bbb... 100%
                                                110 B
pulling 1fa69e2371b7... 100%
                                                154 B
pulling f4c99b0ffe2c... 100%
                                                 32 B
pulling cdf118eda510... 100%
                                                557 B
verifying sha256 digest
writing manifest
success
sunrise@ubuntu:~$
```

3. Use WizardLM-2

3.1. Run WizardLM-2

If the system does not have a running model, the system will automatically pull the WizardLM-2 7B model and run it:

ollama run wizardlm2:7b

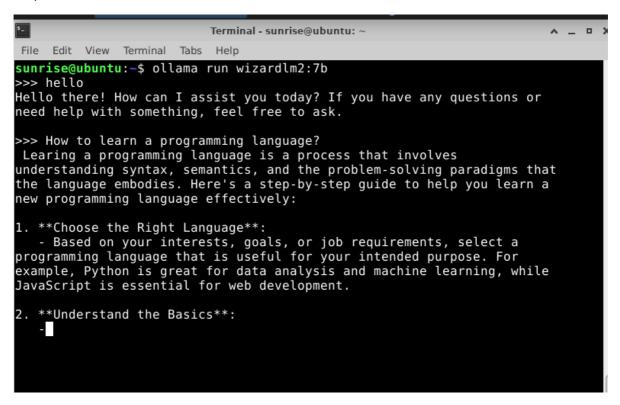


3.2. Have a conversation

How to learn a programming language?

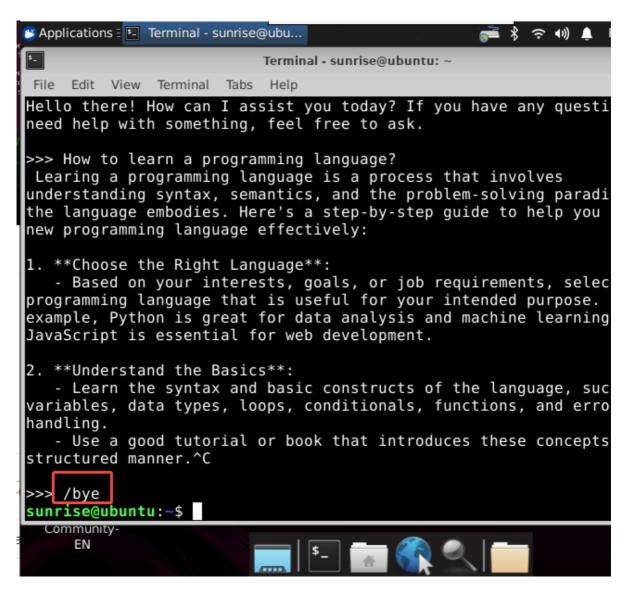
The time required to reply to the question depends on the hardware configuration, please be patient!

Requires **RDK X5 8Gb** board to run.



3.3. End the conversation

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



References

Ollama

Website: https://ollama.com/

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

WizardLM-2

Ollama model: https://ollama.com/library/wizardlm2