

Qwen2 model

Qwen2 model

[Model scale](#)

[Performance performance](#)

[Got Qwen2](#)

[Use Qwen2](#)

[Run Qwen2](#)

[Dialogue](#)

[End conversation](#)

[Reference material](#)

Demonstration environment

Development Board : Raspberry Pi 500

SD(TF)card: 32G (Above 16G, the larger the capacity, the more models can be experienced)

Raspberry Pi 500 (8G RAM): Run 8B and below parameter models

Model scale

Model	Parameter
Qwen2	0.5B
Qwen2	1.5B
Qwen2	7B
Qwen2	72B

Raspberry Pi 500 (8G RAM): Test Qwen2 model with parameters of 7B and below.

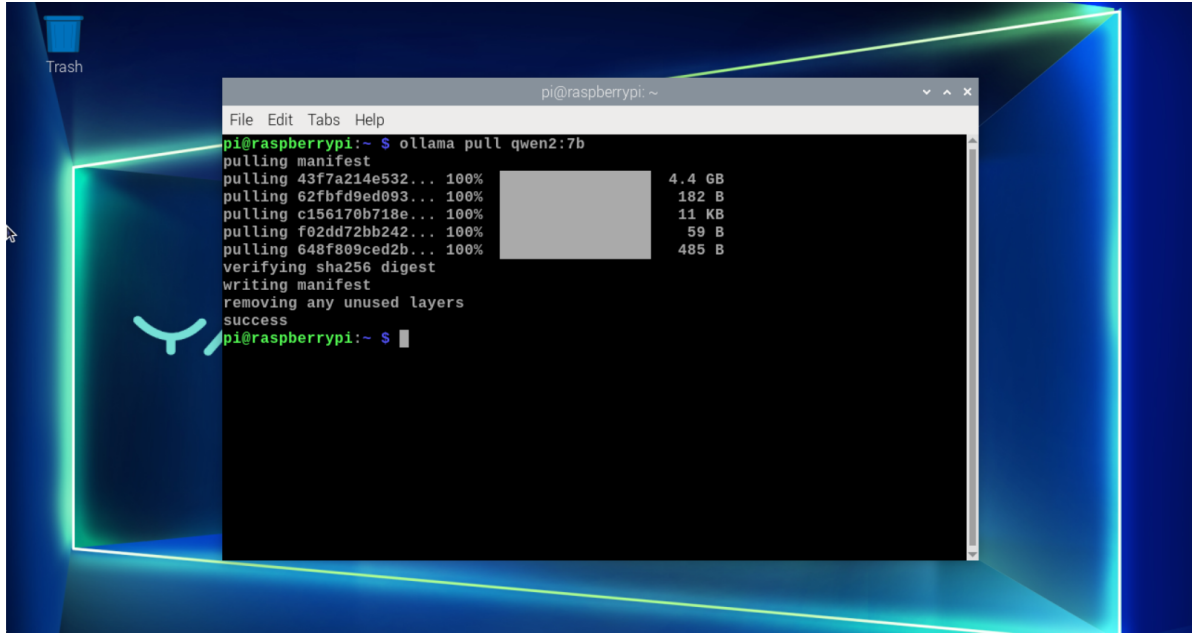
Performance performance

	Qwen2-7B Instruct	Llama3-8B Instruct	GLM4-9B Chat
AlignBench	7.21	6.20	7.01
MT-Bench	8.41	8.05	8.35
MMLU	70.5	68.4	72.4
GSM8K	82.3	79.6	79.6
MATH	49.6	30.0	50.6
HumanEval	79.9	62.2	71.8
C-Eval	77.2	45.9	75.6

Got Qwen2

Using the pull command will automatically pull the models from the Ollama model library: **Due to the memory limitation of SD card, Pi500_AI_Pure image provides qwen2:1.5b model, but does not provide 7b model!**

```
ollama pull qwen2:7b
```



Use Qwen2

Run Qwen2

Raspberry Pi 5B (8G RAM)

If the system does not have a running model, the system will automatically pull the Qwen2 7B model and run it.

```
ollama run qwen2:7b
```

Dialogue

```
Tell me something about large language models.
```

The response time to the question is related to the hardware configuration, please be patient and wait.

```
pi@raspberrypi: ~  
File Edit Tabs Help  
pi@raspberrypi:~ $ ollama run qwen2:7b  
>>> Tell me something about large language models.  
Large language models are sophisticated artificial intelligence systems designed to generate human-like text by learning patterns and structures from vast amounts of textual data. These models have been trained on massive datasets, sometimes containing billions or even trillions of parameters, making them highly complex and capable of producing a wide range of outputs, including essays, stories, poetry, and code.  
  
The training process involves feeding the model large volumes of text, allowing it to learn patterns in language use, semantics, syntax, and context. This enables the model to generate responses that are coherent, relevant, and sometimes even creative.  
  
One key feature of large language models is their ability to handle a variety of input prompts or questions, which they then use to produce output that continues the text in a plausible manner. The more data the model has been trained on, the better it can generalize across different types of content and topics.  
  
These models have numerous applications, including but not limited to:  
  
1. Automated Writing: They are used for generating text, which can be useful for creating content for websites, blogs, or even marketing materials.  
2. Virtual Customer Service: Large language models can simulate human-like conversations with customers, providing personalized responses and potentially improving customer service experiences.  
3. Translation Services: By understanding the nuances of various languages, these models can translate text between languages accurately and efficiently.  
4. Educational Tools: They can be used to create adaptive learning platforms that tailor educational content based on individual student needs.  
5. Research and Development: In fields like medicine, science, and technology, large language models assist in summarizing research papers, generating hypotheses, or even helping with data analysis.  
  
However, they also come with challenges and limitations. Large language models can generate biased or inappropriate content if the training data contains such biases. Additionally, their outputs are not always factually accurate, especially when dealing with new topics that were not part of their training dataset. Furthermore, the computational resources required to train these models are significant and may pose ethical concerns regarding energy consumption.
```

End conversation

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

```
pi@raspberrypi: ~  
File Edit Tabs Help  
pi@raspberrypi:~ $ ollama run qwen2:7b  
>>> Use Ctrl + d or /bye to exit.  
>>>  
pi@raspberrypi:~ $  
  
The training process involves feeding the model large volumes of text, allowing it to learn patterns in language use, semantics, syntax, and context. This enables the model to generate responses that are coherent, relevant, and sometimes even creative.  
  
One key feature of large language models is their ability to handle a variety of input prompts or questions, which they then use to produce output that continues the text in a plausible manner. The more data the model has been trained on, the better it can generalize across different types of content and topics.  
  
These models have numerous applications, including but not limited to:  
  
1. Automated Writing: They are used for generating text, which can be useful for creating content for websites, blogs, or even marketing materials.  
2. Virtual Customer Service: Large language models can simulate human-like conversations with customers, providing personalized responses and potentially improving customer service experiences.  
3. Translation Services: By understanding the nuances of various languages, these models can translate text between languages accurately and efficiently.  
4. Educational Tools: They can be used to create adaptive learning platforms that tailor educational content based on individual student needs.  
5. Research and Development: In fields like medicine, science, and technology, large language models assist in summarizing research papers, generating hypotheses, or even helping with data analysis.  
  
However, they also come with challenges and limitations. Large language models can generate biased or inappropriate content if the training data contains such biases. Additionally, their outputs are not always factually accurate, especially when dealing with new topics that were not part of their training dataset. Furthermore, the computational resources required to train these models are significant and may pose ethical concerns regarding energy consumption.  
  
Overall, large language models represent a powerful tool in the realm of natural language processing (NLP), offering both opportunities for innovation and areas needing careful consideration and management.  
  
>>> /bye  
pi@raspberrypi:~ $
```

Reference material

Ollama

Website: <https://ollama.com/>

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

Qwen2

GitHub: <https://github.com/QwenLM/Qwen2>

Ollama model: <https://ollama.com/library/qwen2>

