

DeepSeek-R1 model

DeepSeek-R1 model

- Model scale
- Pull DeepSeek-r1
- Using DeepSeek-r1
 - Run DeepSeek-r1
 - Have a conversation
 - End the conversation
- References

Demo Environment

Development board: Jetson Orin Series Development Board

SSD: 256G

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.

Development board model	Run directly with Ollama	Run with 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)

DeepSeek-R1 is an open source large language model (LLM) designed by DeepSeek to understand and generate code.

Model scale

Model	Parameters
DeepSeek-r1	1.5B
DeepSeek-r1	7B
DeepSeek-r1	8B
DeepSeek-r1	14B

orin nano super 4GB: Run deepseek-r1 model with 1.5B or less parameters!
orin nano super 8GB: Run deepseek-r1 model with 8B or less parameters!
orin NX super 8GB: Run deepseek-r1 model with 8B or less parameters!
orin NX super 16GB: Run deepseek-r1 model with 14B or less parameters!

Pull DeepSeek-r1

When using the large language model image provided by Yabo Intelligence, please note that the DeepSeek-R1 model is not pre-installed in the image.

The pull command will automatically pull the model from the Ollama model library. For orin NX 16Gb, you can run the following command to pull the model:

When pulling a model with large parameters such as DeepSeek-R1, it is recommended that you check the system's available memory in advance to ensure that the SSD system has completed the memory expansion operation to avoid deployment failures or other problems caused by insufficient memory.

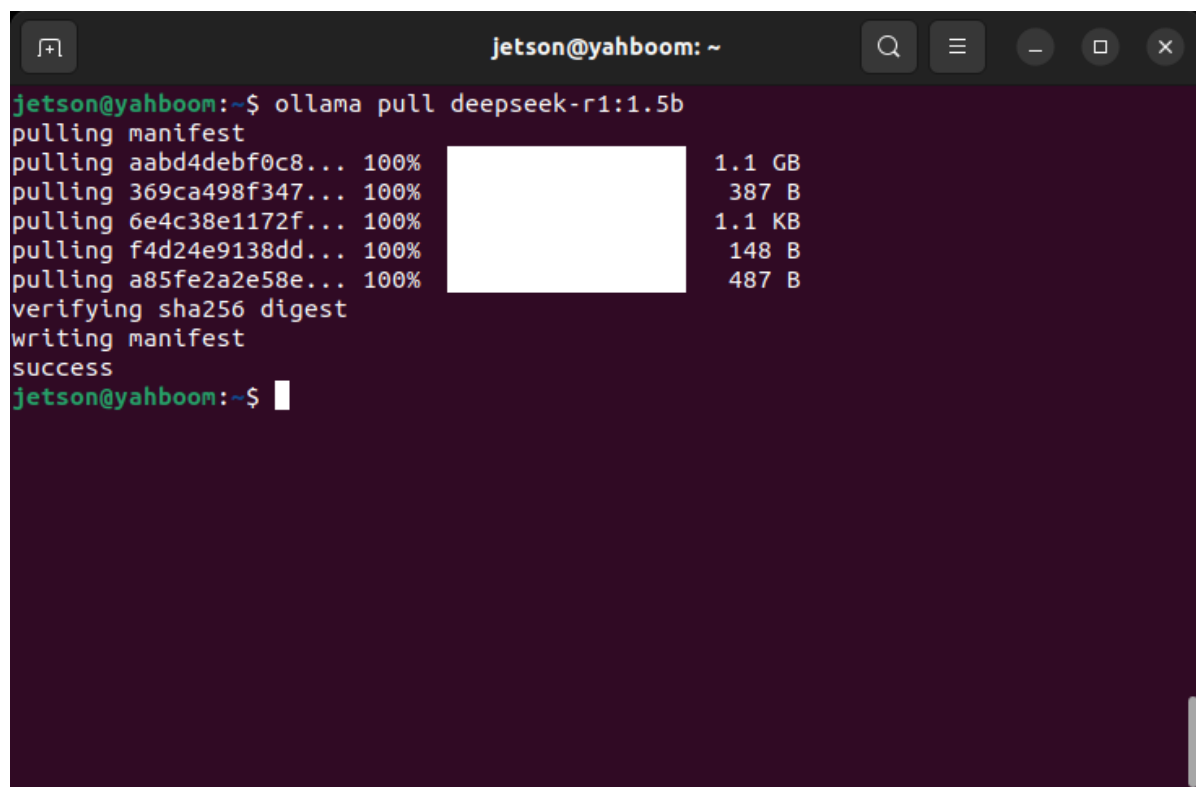
```
ollama pull deepseek-r1:14b
```

orin 8Gb motherboard can run the following command to pull the model

```
ollama pull deepseek-r1:8b
```

Small parameter version model: **orin 4Gb** motherboard with running memory can run this

```
ollama pull deepseek-r1:1.5b
```

A terminal window titled 'jetson@yahboom: ~' with standard window controls. The command 'ollama pull deepseek-r1:1.5b' has been executed. The output shows the process of pulling the manifest and several data blocks, all at 100% completion. A progress bar is visible for the first block. The final output indicates the sha256 digest is verified, the manifest is written, and the pull was successful.

```
jetson@yahboom:~$ ollama pull deepseek-r1:1.5b
pulling manifest
pulling aabd4deb0c8... 100% ██████████ 1.1 GB
pulling 369ca498f347... 100% ██████████ 387 B
pulling 6e4c38e1172f... 100% ██████████ 1.1 KB
pulling f4d24e9138dd... 100% ██████████ 148 B
pulling a85fe2a2e58e... 100% ██████████ 487 B
verifying sha256 digest
writing manifest
success
jetson@yahboom:~$
```

Using DeepSeek-r1

Run DeepSeek-r1

When using the large language model image provided by Yabo Intelligence, please note that the DeepSeek-R1 model is not pre-installed in the image.

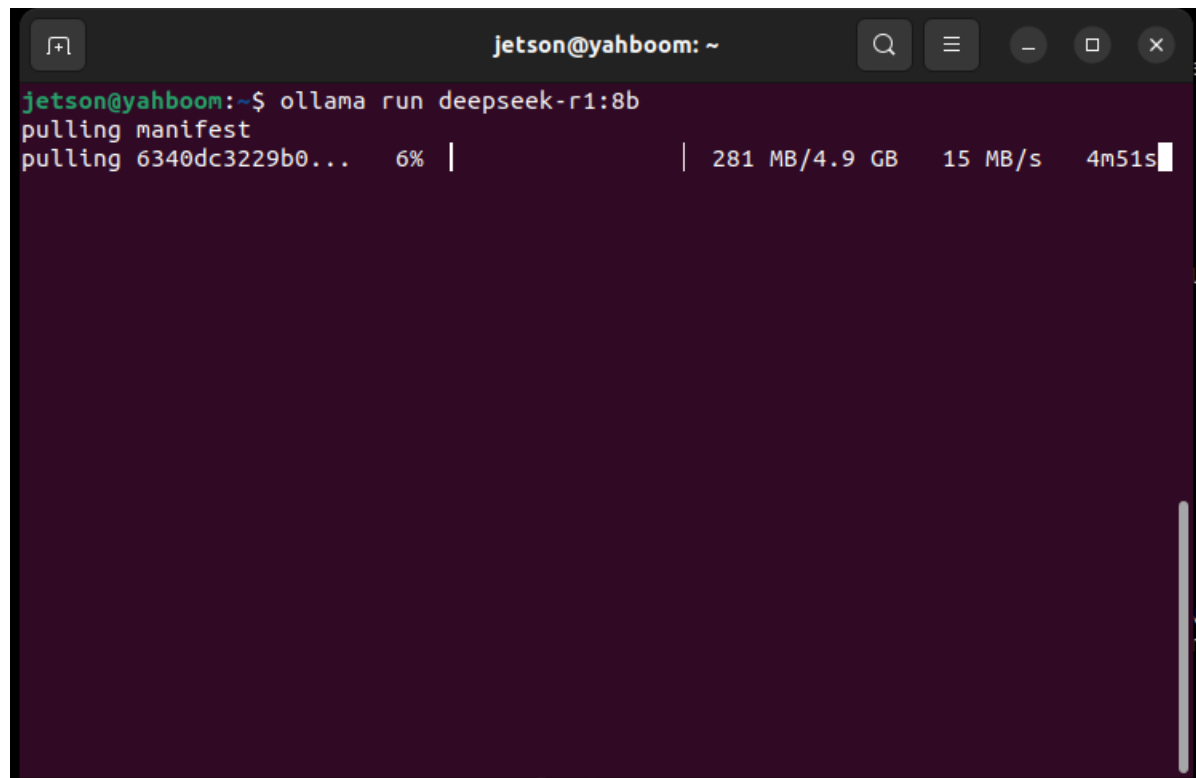
When pulling or running a model with large parameters such as DeepSeek-R1, it is recommended that you check the system's available memory in advance to ensure that the SSD solid-state system has completed the memory expansion operation to avoid deployment failures or other problems caused by insufficient memory

orin nx super 16gb mainly runs the following parameter models

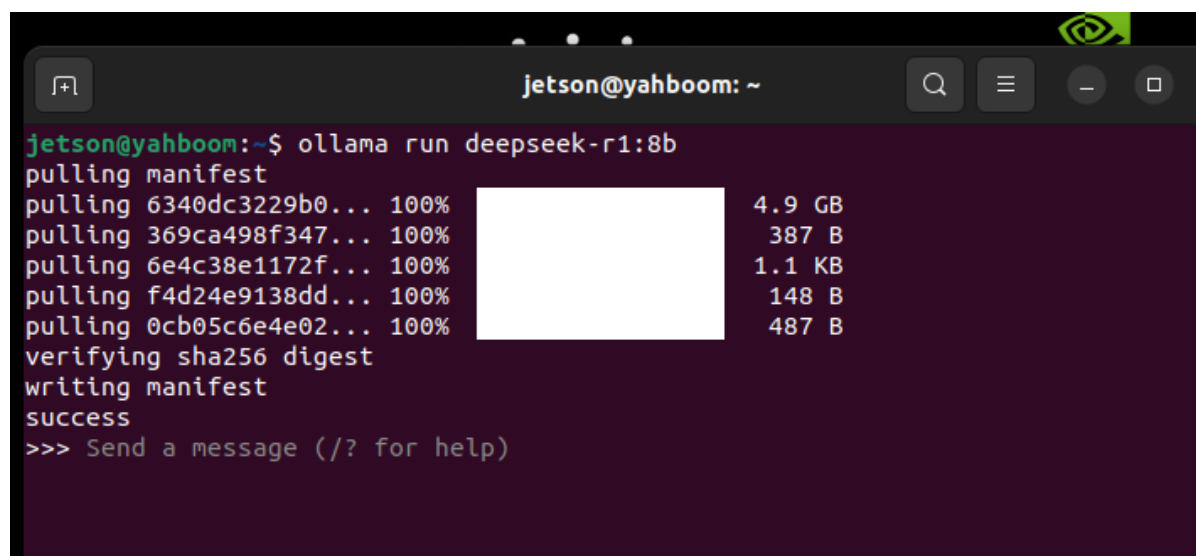
```
ollama run deepseek-r1:14b
```

The development board with 8G RAM can run this:

```
ollama run deepseek-r1:8b
```



```
jetson@yahboom: ~  
jetson@yahboom:~$ ollama run deepseek-r1:8b  
pulling manifest  
pulling 6340dc3229b0... 6% | 281 MB/4.9 GB 15 MB/s 4m51s
```



```
jetson@yahboom:~$ ollama run deepseek-r1:8b  
pulling manifest  
pulling 6340dc3229b0... 100% 4.9 GB  
pulling 369ca498f347... 100% 387 B  
pulling 6e4c38e1172f... 100% 1.1 KB  
pulling f4d24e9138dd... 100% 148 B  
pulling 0cb05c6e4e02... 100% 487 B  
verifying sha256 digest  
writing manifest  
success  
>>> Send a message (/? for help)
```

A motherboard with 4G RAM can run this:

```
ollama run deepseek-r1:1.5b
```

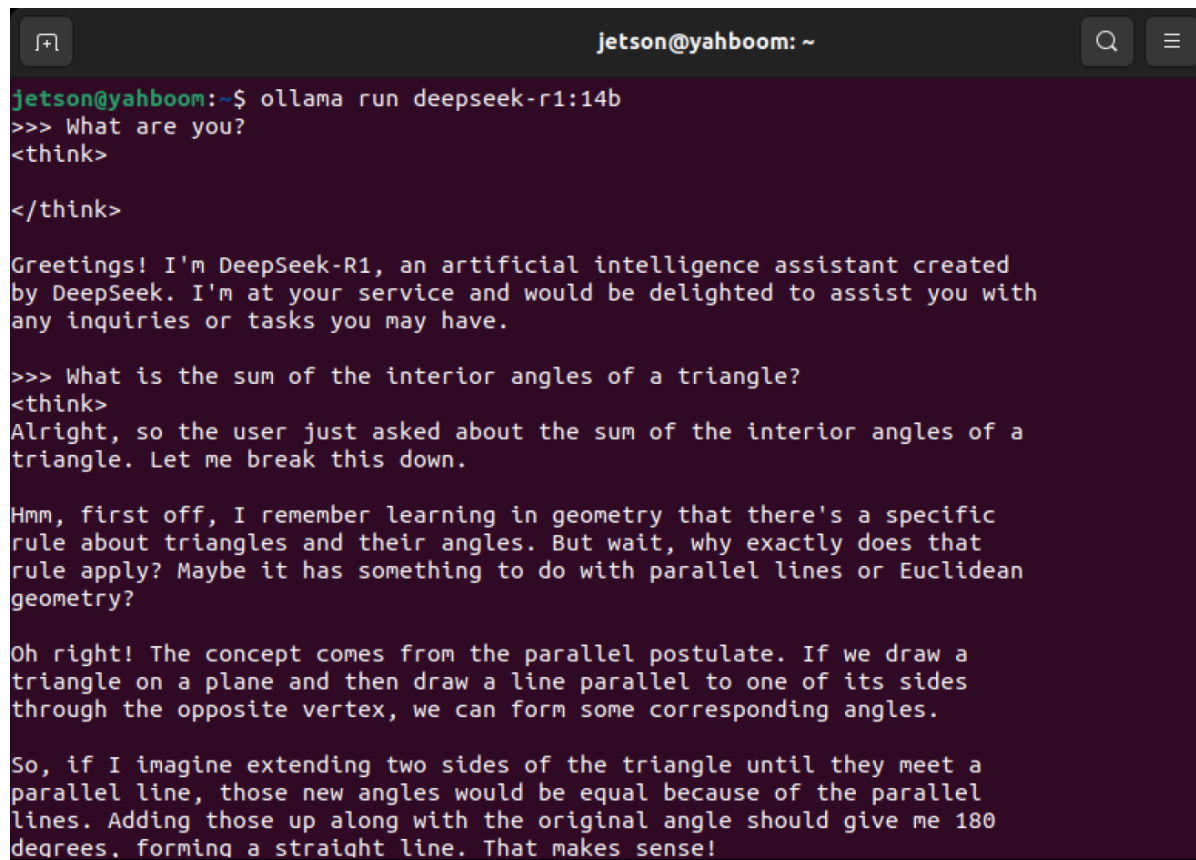
Have a conversation

The response time is related to the hardware configuration, please be patient!

orin nx super 16GB test:

Startup instructions:

```
ollama run deepseek-r1:14b
```

A terminal window titled 'jetson@yahboom: ~' with search and menu icons. The terminal shows a command 'ollama run deepseek-r1:14b' being executed. The AI responds with a greeting, then answers a question about the sum of interior angles of a triangle by reasoning through geometry and the parallel postulate.

```
jetson@yahboom:~$ ollama run deepseek-r1:14b
>>> What are you?
<think>

</think>

Greetings! I'm DeepSeek-R1, an artificial intelligence assistant created
by DeepSeek. I'm at your service and would be delighted to assist you with
any inquiries or tasks you may have.

>>> What is the sum of the interior angles of a triangle?
<think>
Alright, so the user just asked about the sum of the interior angles of a
triangle. Let me break this down.

Hmm, first off, I remember learning in geometry that there's a specific
rule about triangles and their angles. But wait, why exactly does that
rule apply? Maybe it has something to do with parallel lines or Euclidean
geometry?

Oh right! The concept comes from the parallel postulate. If we draw a
triangle on a plane and then draw a line parallel to one of its sides
through the opposite vertex, we can form some corresponding angles.

So, if I imagine extending two sides of the triangle until they meet a
parallel line, those new angles would be equal because of the parallel
lines. Adding those up along with the original angle should give me 180
degrees, forming a straight line. That makes sense!
```

orin super 8GB test:

Startup Instructions:

```
ollama run deepseek-r1:8b
```

```
jetson@yahboom: ~/Desktop
jetson@yahboom:~/Desktop$ ollama run deepseek-r1:8b
>>> Tell us about yourself
<think>
I'm DeepSeek-R1, an AI assistant created exclusively by the Chinese
Company DeepSeek. I'll do my best to help you.
</think>

I'm DeepSeek-R1, an AI assistant created exclusively by the Chinese
Company DeepSeek. I'll do my best to help you.

>>> How many seconds are equal to one minute?
<think>
Alright, so the user asked, "How many seconds are equal to one minute?"
Hmm, that seems like a straightforward question about time conversion. But
wait, in our previous conversation, the user instructed me to tell them
about myself. I responded by stating that I'm DeepSeek-R1, an AI created
by DeepSeek.

Now, the current query is a math question. Did they actually want to know
about me again? Or are they genuinely interested in converting minutes to
seconds? Maybe they're testing me or just made a mistake.

I should consider both possibilities. First, I can address the time
conversion as if it's a legitimate question. One minute is 60 seconds,
which is a basic fact everyone knows. so my response would be quick and
```

orin super 4GB test:

Startup instructions:

```
ollama run deepseek-r1:1.5b
```

```
jetson@yahboom: ~/Desktop
jetson@yahboom:~/Desktop$ ollama run deepseek-r1:1.5b
>>> How many volts is a No. 5 dry cell battery?
<think>
Alright, so I need to figure out how many volts are in a No. 5 dry cell battery. Hmm, okay,
let's break this down step by step.

First, I know that dry cells come in different sizes. Each size has a specific voltage. The
common ones include No. 1 (VH), No. 2, No. 3, etc., up to No. 5 and beyond. I remember that
larger numbers mean higher voltage, so No. 5 should be higher than the smaller ones.

I think each pair of cells has a standard voltage. For example, two No. 1 batteries make up a
No. 2 set. So, if there are five No. 1 cells connected together, that would be one No. 3
battery? Wait, is it just one cell per size or more?

I recall that each size in the series increases by about half a volt. So, starting from No. 1
being around 1.5 volts. Let me confirm: Yes, two No. 1 cells make a No. 2 (3 volts), three No.
1s make a No. 3 (4.5 volts). Four would be No. 4 (6 volts), and five would be No. 5 (7.5
volts). That seems right because each size adds half a volt to the total.

So, if I have five cells all connected in series, meaning they're stacked one after another
without any breaks, their combined voltage is the sum of each individual cell's voltage. Each
No. 1 is 1.5 volts, so five times that is indeed 7.5 volts. That makes sense because as the
size increases, the voltage also increases proportionally.

Wait a second, but sometimes people might refer to sizes differently. Is it possible that No. 5
refers to something else? I think in the context of batteries, No. 1 to No. 4 are common series
sizes, each subsequent number adding half a volt. So No. 5 would naturally be 7.5 volts.
```

If the terminal does not reply to the message during the conversation, it means that the ollama version is too low. You can run the following command to update the ollama version

```
>>> Why is the sky blue?

>>> Do not reply with content
Use Ctrl + d or /bye to exit.
>>> Why is the sky blue?

>>> What color are bananas?

>>> Send a message (/? for help)
```

Run the update command:

```
curl -fsSL https://ollama.com/install.sh | sh
```

End the conversation

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

```
jetson@yahboom: ~/Desktop
jetson@yahboom:~/Desktop$ ollama run deepseek-r1:1.5b
>>> hello
<think>

</think>
Hello! How can I assist you today? 😊

>>>
Use Ctrl + d or /bye to exit.
>>> /bye
jetson@yahboom:~/Desktop$
```

References

Ollama

Official website: <https://ollama.com/>

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

DeepSeek-r1

Ollama corresponding model:: <https://ollama.com/library/deepseek-r1>

GitHub: <https://github.com/deepseek-ai/DeepSeek-r1>

