# Phi-3 model

#### Phi-3 model

Model scale
Performance performance
Got Phi-3
Use Phi-3
Run Phi-3
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

Phi-3 is a powerful and cost-effective Small Language Model (SLM) launched by Microsoft, which outperforms models of the same size and higher in various language, inference, coding, and mathematical benchmarks.

## **Model scale**

| Model          | Parameter |  |  |  |  |
|----------------|-----------|--|--|--|--|
| Phi-3 (Mini)   | 3.8B      |  |  |  |  |
| Phi-3 (Medium) | 14B       |  |  |  |  |

Raspberry Pi 500 (8G RAM): Test Phi-3 model with 3.8B parameters!

## **Performance performance**

| Category                           | Benchmark                             | Phi-3                |                        |                          |                           |          |            |              | Hama 2            | CDT2 F                | Claud- 3           |
|------------------------------------|---------------------------------------|----------------------|------------------------|--------------------------|---------------------------|----------|------------|--------------|-------------------|-----------------------|--------------------|
|                                    |                                       | Phi-3-<br>Mini-4K-In | Phi-3-Mini-<br>128K-In | Phi-3-Small<br>(Preview) | Phi-3-Medium<br>(Preview) | Gemma-7b | Mistral-7b | Mixtral-8x7b | Llama-3-<br>8B-In | GPT3.5-<br>Turbo-1106 | Claude-3<br>Sonnet |
| Popular<br>Aggregate<br>Benchmarks | AGI Eval<br>(0-shot)                  | 37.5                 | 36.9                   | 45                       | 48.4                      | 42.1     | 35.1       | 45.2         | 42                | 48.4                  | 48.4               |
|                                    | MMLU<br>(5-shot)                      | 68.8                 | 68.1                   | 75.6                     | 78.2                      | 63.6     | 61.7       | 70.5         | 66.5              | 71.4                  | 73.9               |
|                                    | BigBench Hard<br>(0-shot)             | 71.7                 | 71.5                   | 74.9                     | 81.3                      | 59.6     | 57.3       | 69.7         | 51.5              | 68.3                  |                    |
| Language<br>Understanding          | ANLI<br>(7-shot)                      | 52.8                 | 52.8                   | 55                       | 58.7                      | 48.7     | 47.1       | 55.2         | 57.3              | 58.1                  | 68.6               |
|                                    | HellaSwag<br>(5-shot)                 | 76.7                 | 74.5                   | 78.7                     | 83                        | 49.8     | 58.5       | 70.4         | 71.1              | 78.8                  | 79.2               |
| Reasoning                          | ARC Challenge<br>(10-shot)            | 84.9                 | 84                     | 90.7                     | 91                        | 78.3     | 78.6       | 87.3         | 82.8              | 87.4                  | 91.6               |
|                                    | ARC Easy<br>(10-shot)                 | 94.6                 | 95.2                   | 97.1                     | 97.8                      | 91,4     | 90.6       | 95.6         | 93.4              | 96.3                  | 97.7               |
|                                    | BoolQ<br>(0-shot)                     | 77.6                 | 78.7                   | 82.9                     | 86.6                      | 66       | 72.2       | 76.6         | 80.9              | 79.1                  | 87.1               |
|                                    | CommonsenseQA<br>(10-shot)            | 80.2                 | 78                     | 80.3                     | 82.6                      | 76.2     | 72.6       | 78.1         | 79                | 79.6                  | 82.6               |
|                                    | MedQA<br>(2-shot)                     | 53.8                 | 55.3                   | 58.2                     | 69.4                      | 49.6     | 50         | 62.2         | 60.5              | 63.4                  | 67.9               |
|                                    | OpenBookQA<br>(10-shot)               | 83.2                 | 80.6                   | 88.4                     | 87.2                      | 78.6     | 79.8       | 85.8         | 82.6              | 86                    | 90.8               |
|                                    | PIQA<br>(S-shot)                      | 84.2                 | 83.6                   | 87.8                     | 87.7                      | 78.1     | 77.7       | 86           | 75.7              | 86.6                  | 87.8               |
|                                    | Social IQA<br>(5-shot)                | 76.6                 | 76.1                   | 79                       | 80.2                      | 65.5     | 74.6       | 75.9         | 73.9              | 68.3                  | 80.2               |
|                                    | TruthfulQA (MC2)<br>(10-shot)         | 65                   | 63.2                   | 68.7                     | 75.7                      | 52.1     | 53         | 60.1         | 63.2              | 67.7                  | 77.8               |
|                                    | WinoGrande<br>(5-shot)                | 70.8                 | 72.5                   | 82.5                     | 81.4                      | 55.6     | 54.2       | 62           | 65                | 68.8                  | 81.4               |
| Factual<br>Knowledge               | TriviaQA<br>(5-shot)                  | 64                   | 57.1                   | 59.1                     | 75.6                      | 72.3     | 75.2       | 82.2         | 67.7              | 85.8                  | 65.7               |
| Math                               | GSM8K Chain of<br>Thought<br>(0-shot) | 82.5                 | 83.6                   | 88.9                     | 90.3                      | 59.8     | 46.4       | 64.7         | 77.4              | 78.1                  | 79.1               |
| Code<br>generation                 | HumanEval<br>(0-shot)                 | 59.1                 | 57.9                   | 59.1                     | 55.5                      | 34.1     | 28         | 37.8         | 60.4              | 62.2                  | 65.9               |
|                                    | MBPP<br>(3-shot)                      | 53.8                 | 62.5                   | 71.4                     | 74.5                      | 51.5     | 50.8       | 60.2         | 67.7              | 77.8                  | 79,4               |

## **Got Phi-3**

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

ollama pull phi3:3.8b

```
File Edit Tabs Help

ph@raspberrypi: ~ $ ollama pull phi3:3.8b

pulling pace6713dc74... 100%

pulling $2823554846f... 100%

pulling $4282554846f... 100%

pulling $4282547179... 100%

pulling $42825417179... 100%

pulling $4285417179... 100%

pulling $42854177179... 100%

pulling $42854177179... 100%

pulling $42854177179... 100%

pulling $42854177179... 100%

pulling $42854177179...
```

Use Phi-3

#### Run Phi-3

If the system does not have a running model, the system will automatically obtain the Phi-3 3.8B model and run it: **The Pi500\_AI\_Pure image does not provide this model!** 

```
ollama run phi3:3.8b
```

## **Dialogue**

```
Write a joke about dogs

Change another one

Tell me a bedtime story
```

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

```
pi@raspberrypi:~ S ollama run phi3:3.8b

>>> Write a joke about dogs
Why don't dogs make good comedians?

Because their jokes always come out "woof"!

>>> Change another one
Alright, here's a slightly more complex dog-related joke:
Why was the math book sad at the dog park?

Because it had too many problems and not enough friends!

>>> Tell me a bedtime story
Once upon a time in a faraway galaxy, there lived a twinkling star named Stella. She shone brightly amongst her celestial family, providing light to the countless planets orbiting around her home. However, despite being loved by all, she felt something was missing from her life - companionship.

One day, while gazing at Earth through a telescope, Stella noticed tiny creatures moving on the blue-green planet below. These were children, playing and laughing under the moonlight. One particular child caught Stella's attention - Lily, who loved to look up at the stars every night before bedtime, dreaming of adventures in space.

Stella felt a strong connection with Lily; her twinkling light was like an invitation for friendship. So she decided to send a special message down through the vast cosmos - a shooting star that landed on Earth right in front of Lily's bedroom window!

Lily, filled with wonder and excitement, ran out into the night sky holding her father's camera. As they captured images of Stella's radiant tail, she noticed something magical unfolding before her eyes - a secret garden hidden within Earth's atmosphere. The plants were glowing, each one representing different emotions: joyful sunflowers, serene lavender, and even tiny pumpkins with witty sayings!

Lily soon discovered that the garden was tended to by Stella and her celestial friends - comets, planets, moons, and calaxies. They had formed a mystical community, each helping others express their emotions through these unique glowing
```

## **End conversation**

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

```
File Edit Tabs Help

celestial family, providing light to the countless planets orbiting around her home. However, despite being loved by
all, she felt something was missing from her life - companionship.

One day, while gazing at Earth through a telescope, Stella noticed tiny creatures moving on the blue-green planet below. These were children, playing and laughing under the moonlight one particular child caught Stella's attention - Lily, who loved to look up at the stars every night before boddien, decaping of educatures in Jones (Stella felt a strong connection with Lily; her send a special message down through the vast of bedroom window!

Lily, filled with wonder and excitement, ran oluse Ctrl + d or /bye to exit.

of Stella's radiant tail, she noticed somethin >>> Earth's atmosphere. The plants were glowing, epigraspberrypi:- $ lavender, and even tiny pumpkins with witty sa

Lily soon discovered that the garden was tender galaxies. They had formed a mystical community plants. The garden provided not only a beautiff

Thrilled with the discovery, Lily spent her niftlends. As she grew older, Stella's light bec. friendships, and nurture compassion towards al

And so, the bond between Earth and Stella cont. mission to spread kindness throughout the gala: stars.

The End.

>>> /bye
pi@raspberrypi:- $ []
```

### Reference material

Ollama

Website: https://ollama.com/

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

Phi-3

Ollama corresponding model: <a href="https://ollama.com/library/phi3">https://ollama.com/library/phi3</a>