


```
!pip install transformers torch
Requirement already satisfied: Cuda11.7.1 in /usr/local/lib/python3.11/dist-packages (from requests / transformers) (2020.1.01)
Downloading nvidia_cublas_cul2-12.4.5.8-py3-none-manylinux2014_x86_64.whl (363.4 MB)
363.4/363.4 MB 2.9 MB/s eta 0:00:00
Downloading nvidia_cuda_cupti_cul2-12.4.127-py3-none-manylinux2014_x86_64.whl (13.8 MB)
13.8/13.8 MB 36.3 MB/s eta 0:00:00
Downloading nvidia_cuda_nvrtc_cul2-12.4.127-py3-none-manylinux2014_x86_64.whl (24.6 MB)
24.6/24.6 MB 15.3 MB/s eta 0:00:00
Downloading nvidia_cuda_runtime_cul2-12.4.127-py3-none-manylinux2014_x86_64.whl (883 kB)
883.7/883.7 kB 35.2 MB/s eta 0:00:00
Downloading nvidia_cudnn_cul2-9.1.0.70-py3-none-manylinux2014_x86_64.whl (664.8 MB)
664.8/664.8 MB 1.2 MB/s eta 0:00:00
Downloading nvidia_cufft_cul2-11.2.1.3-py3-none-manylinux2014_x86_64.whl (211.5 MB)
211.5/211.5 MB 5.8 MB/s eta 0:00:00
Downloading nvidia_curand_cul2-10.3.5.147-py3-none-manylinux2014_x86_64.whl (56.3 MB)
56.3/56.3 MB 10.4 MB/s eta 0:00:00
Downloading nvidia_cusolver_cul2-11.6.1.9-py3-none-manylinux2014_x86_64.whl (127.9 MB)
127.9/127.9 MB 6.8 MB/s eta 0:00:00
Downloading nvidia_cusparses_cul2-12.3.1.170-py3-none-manylinux2014_x86_64.whl (207.5 MB)
207.5/207.5 MB 6.8 MB/s eta 0:00:00
Downloading nvidia_nvjitlink_cul2-12.4.127-py3-none-manylinux2014_x86_64.whl (21.1 MB)
21.1/21.1 MB 64.5 MB/s eta 0:00:00
Installing collected packages: nvidia-nvjitlink-cul2, nvidia-curand-cul2, nvidia-cufft-cul2, nvidia-cuda-runtime-cul2, nvidia-cuda-nvrtc-cul2,
Attempting uninstall: nvidia-nvjitlink-cul2
Found existing installation: nvidia-nvjitlink-cul2 12.5.82
Uninstalling nvidia-nvjitlink-cul2-12.5.82:
Successfully uninstalled nvidia-nvjitlink-cul2-12.5.82
Attempting uninstall: nvidia-curand-cul2
Found existing installation: nvidia-curand-cul2 10.3.6.82
Uninstalling nvidia-curand-cul2-10.3.6.82:
Successfully uninstalled nvidia-curand-cul2-10.3.6.82
Attempting uninstall: nvidia-cufft-cul2
Found existing installation: nvidia-cufft-cul2 11.2.3.61
Uninstalling nvidia-cufft-cul2-11.2.3.61:
Successfully uninstalled nvidia-cufft-cul2-11.2.3.61
Attempting uninstall: nvidia-cuda-runtime-cul2
Found existing installation: nvidia-cuda-runtime-cul2 12.5.82
Uninstalling nvidia-cuda-runtime-cul2-12.5.82:
Successfully uninstalled nvidia-cuda-runtime-cul2-12.5.82
Attempting uninstall: nvidia-cuda-nvrtc-cul2
Found existing installation: nvidia-cuda-nvrtc-cul2 12.5.82
Uninstalling nvidia-cuda-nvrtc-cul2-12.5.82:
Successfully uninstalled nvidia-cuda-nvrtc-cul2-12.5.82
Attempting uninstall: nvidia-cuda-cupti-cul2
Found existing installation: nvidia-cuda-cupti-cul2 12.5.82
Uninstalling nvidia-cuda-cupti-cul2-12.5.82:
Successfully uninstalled nvidia-cuda-cupti-cul2-12.5.82
Attempting uninstall: nvidia-cublas-cul2
Found existing installation: nvidia-cublas-cul2 12.5.3.2
Uninstalling nvidia-cublas-cul2-12.5.3.2:
Successfully uninstalled nvidia-cublas-cul2-12.5.3.2
Attempting uninstall: nvidia-cusparses-cul2
Found existing installation: nvidia-cusparses-cul2 12.5.1.3
Uninstalling nvidia-cusparses-cul2-12.5.1.3:
Successfully uninstalled nvidia-cusparses-cul2-12.5.1.3
Attempting uninstall: nvidia-cudnn-cul2
Found existing installation: nvidia-cudnn-cul2 9.3.0.75
Uninstalling nvidia-cudnn-cul2-9.3.0.75:
Successfully uninstalled nvidia-cudnn-cul2-9.3.0.75
```

```
from transformers import pipeline

generator = pipeline("text-generation", model="deepseek-ai/DeepSeek-R1-Distill-Qwen-1.5B", device_map="auto")

result=generator(
    "一個農場有雞跟兔，共有5個頭16隻腳，請問雞跟兔各有幾隻?請用繁體中文回答",
    truncation=True,
    max_length=30000,
    num_return_sequences=1,
)

print(result)
```

 config.json: 100% 679/679 [00:00<00:00, 49.7kB/s]

model.safetensors: 100% 3.55G/3.55G [00:43<00:00, 96.2MB/s]

Sliding Window Attention is enabled but not implemented for `sdpa`; unexpected results may be encountered.

generation_config.json: 100% 181/181 [00:00<00:00, 6.05kB/s]

tokenizer_config.json: 100% 3.07k/3.07k [00:00<00:00, 125kB/s]

tokenizer.json: 100% 7.03M/7.03M [00:00<00:00, 14.0MB/s]

Device set to use cpu

[{'generated_text': '一個農場有雞跟兔，共有5個頭16隻腳，請問雞跟兔各有幾隻?請用繁體中文回答。\\n\\n好的，我现在来解决这个问题。题目说一个农场有'}]