

DEEPSEEK – R1

Ollama(deepseek-r1) + Langchain + Streamlit

- ✓ Min system Requirement : 8gb Ram
- ✓ Login to → <https://ollama.com/> → Download & install Ollama →

The image shows two screenshots from a web browser. The top screenshot is the Ollama homepage, which features a search bar, links to Discord, GitHub, and Models, and a 'Download' button. The bottom screenshot is the 'deepseek-r1' model page, which includes a description of the model, a list of available sizes (1.5b, 7b, 8b, 14b, 32b, 70b, 671b), and a table of model details.

deepseek-r1

DeepSeek's first-generation of reasoning models with comparable performance to OpenAI-o1, including six dense models distilled from DeepSeek-R1 based on Llama and Qwen.

1.5b 7b 8b 14b 32b 70b 671b

6.4M Pulls Updated 11 days ago

7b 28 Tags ollama run deepseek-r1

Updated 12 days ago	0a8c26691023 · 4.7GB
model	arch qwen2 · parameters 7.62B · quantization Q4_K_M 4.7GB
params	{ "stop": ["< begin_of_sentence >", "< end_of_sentence >"], 148B
template	{{- if .System }}{{ .System }}{{ end }} {{- range \$i, \$_ := .Mes... 387B
license	MIT License Copyright (c) 2023 DeepSeek Permission is hereby gra... 1.1kB

- ✓ **DeepSeek-R1 model locally.** The system requirements vary significantly based on model size.
 - **1.5B model:** Works fine on 8GB RAM and 4GB+ VRAM (even integrated GPUs can handle it with CPU offloading).
 - **7B - 8B** models: Need 16GB+ RAM and 12GB+ VRAM (RTX 3060 or better) for smooth performance.
 - **14B+ models:** Require at least 24GB VRAM (RTX 3090/4090) and 32GB+ RAM for decent speed.
 - **32B - 70B models:** Need A100 80GB+ or multiple GPUs, plus 64GB+ RAM (offloading works but is slow).

- **671B model** : Impossible to run locally—it requires huge cloud compute resources.

🖥️ System Cmd → Ollama →

```
C:\Users\A3MAX SOFTWARE TECH>ollama
Usage:
  ollama [flags]
  ollama [command]

Available Commands:
  serve      Start ollama
  create      Create a model from a Modelfile
  show        Show information for a model
  run         Run a model
  stop        Stop a running model
  pull        Pull a model from a registry
  push        Push a model to a registry
  list        List models
  ps          List running models
  cp          Copy a model
  rm          Remove a model
  help        Help about any command

Flags:
  -h, --help      help for ollama
  -v, --version    Show version information

Use "ollama [command] --help" for more information about a command.

C:\Users\A3MAX SOFTWARE TECH>
```

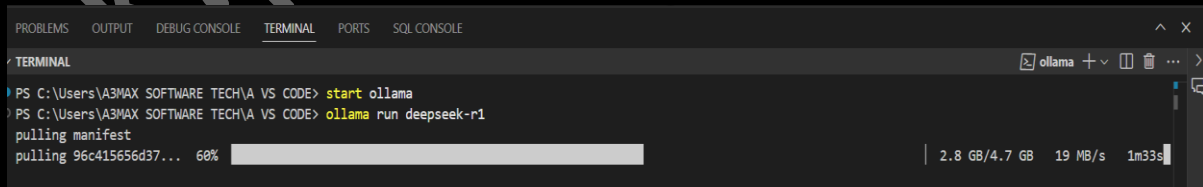
🖥️ **Vscode** → terminal → Ollama run deepseek-r1 → first it will download 4.7GB & then it will give a prompt to use user int

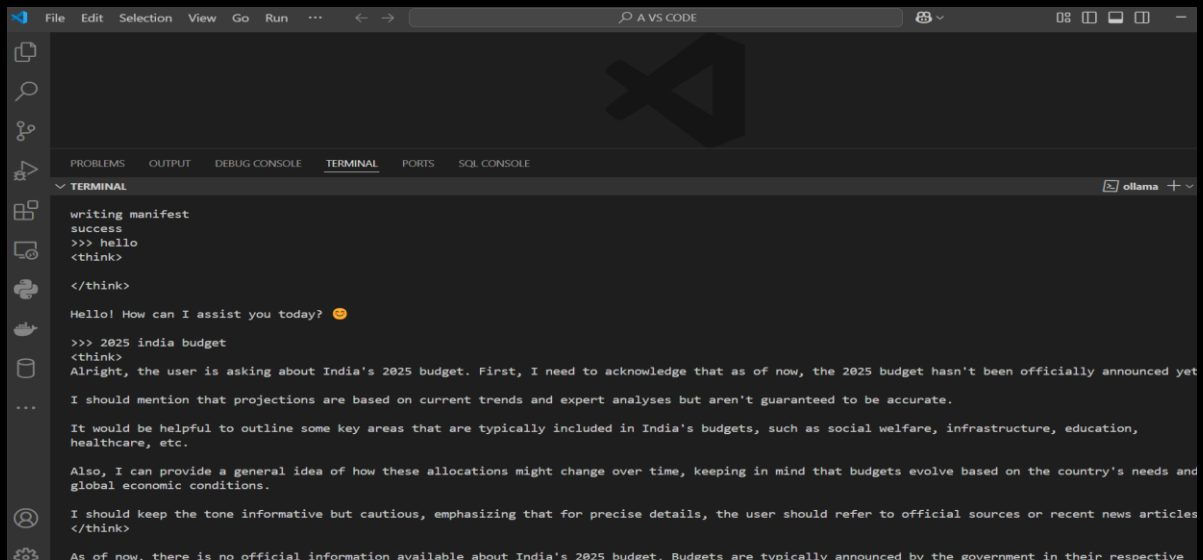
```
C:\Users\A3MAX SOFTWARE TECH>ollama model list
Error: unknown command "model" for "ollama"

C:\Users\A3MAX SOFTWARE TECH>ollama list
NAME                ID                SIZE      MODIFIED
deepseek-r1:latest  0a8c26691023      4.7 GB    5 days ago
llama2:latest       78e26419b446      3.8 GB    7 days ago

C:\Users\A3MAX SOFTWARE TECH>ollama pull deepseek-r1
pulling manifest
pulling 96c415656d37... 100%
pulling 369ca498f347... 100%
pulling 6e4c38e1172f... 100%
pulling f4d24e9138dd... 100%
pulling 40fb844194b2... 100%
verifying sha256 digest
writing manifest
success
```

```
C:\Users\A3MAX SOFTWARE TECH>ollama run deepseek-r1
>>> Send a message (/? for help)
```





```
File Edit Selection View Go Run ... < -> A VS CODE
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS SQL CONSOLE
TERMINAL ollama +
writing manifest
success
>>> hello
<think>
</think>
Hello! How can I assist you today? 🌟
>>> 2025 india budget
<think>
Alright, the user is asking about India's 2025 budget. First, I need to acknowledge that as of now, the 2025 budget hasn't been officially announced yet.
I should mention that projections are based on current trends and expert analyses but aren't guaranteed to be accurate.
It would be helpful to outline some key areas that are typically included in India's budgets, such as social welfare, infrastructure, education, healthcare, etc.
Also, I can provide a general idea of how these allocations might change over time, keeping in mind that budgets evolve based on the country's needs and global economic conditions.
I should keep the tone informative but cautious, emphasizing that for precise details, the user should refer to official sources or recent news articles.
</think>
As of now, there is no official information available about India's 2025 budget. Budgets are typically announced by the government in their respective
```

🧩 Let's integrate OLLAMA- DEEPSEEK along with LANGCHAIN FRAMEWORK

🧩 VS Code → Create new folder → Ollama → Install below library

####Install packages

pip install --upgrade langchain langchain-community

Provides access to the ollama models.

pip install -U langchain-ollama

Install streamlit.

pip install streamlit

➤ Create new python file & start write the code