## 本地部署deepseek-r1

## 系统信息：

GPU : NVIDIA GeForce RTX 4050 Laptop GPU

VRAM : 6.44 GB

CPU : AMD64 Family 25 Model 68 Stepping 1, AuthenticAMD

RAM : 16.37 GB

SYSTEM ：win11

## 部署

1：先进入奥拉马官网下载奥拉马

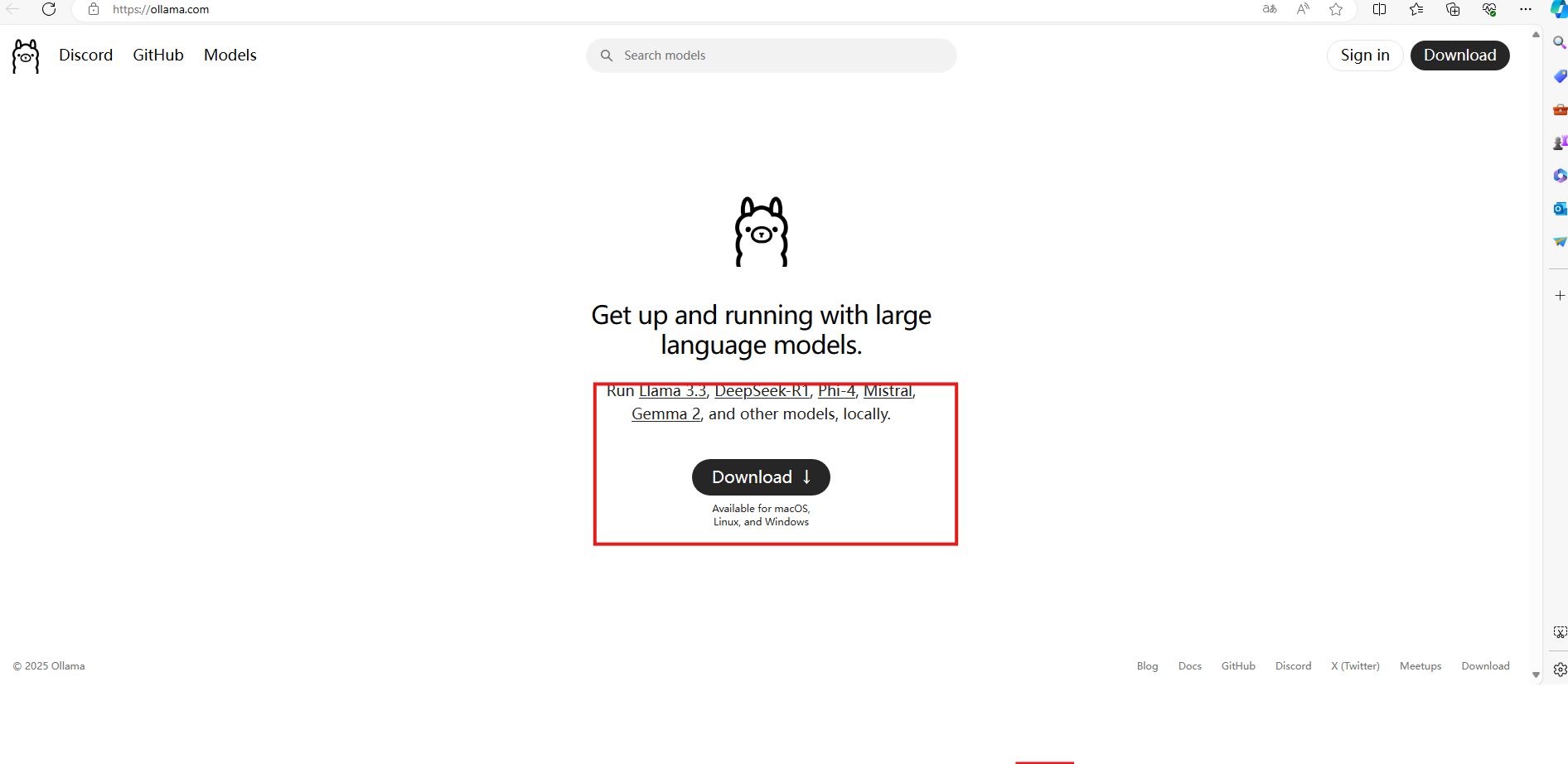
官网地址：

<https://ollama.com/>（[奥拉马](https://ollama.com/)）

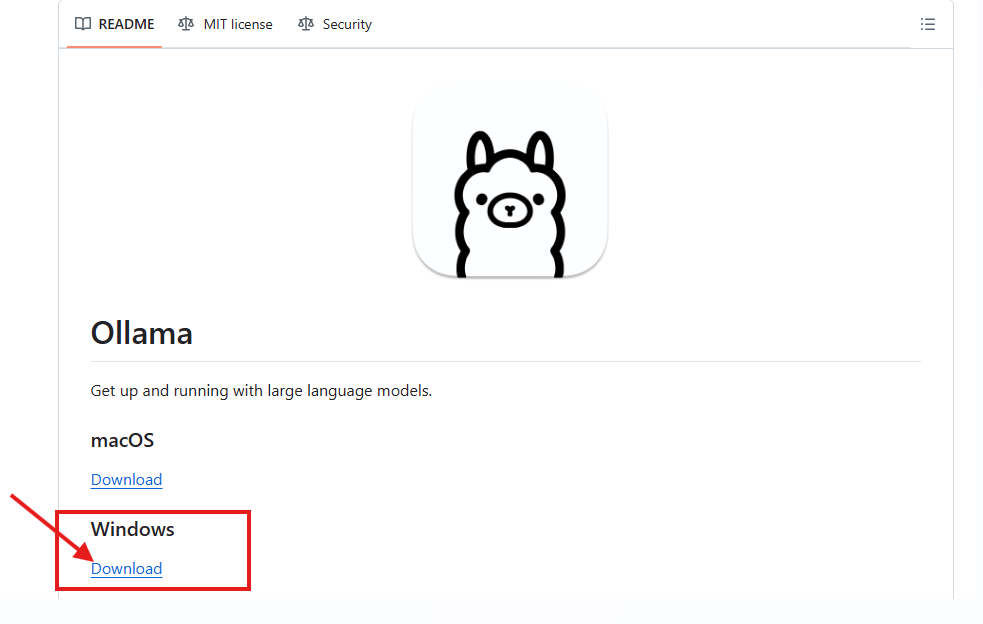
Github地址：

[ollama/ollama：启动并运行 Llama 3.3、DeepSeek-R1、Phi-4、Gemma 2 和其他大型语言模型。](https://github.com/ollama/ollama)

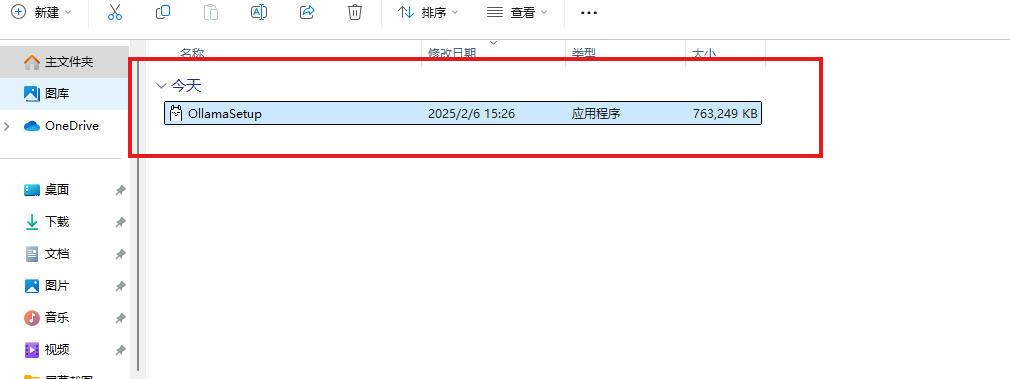
官网：

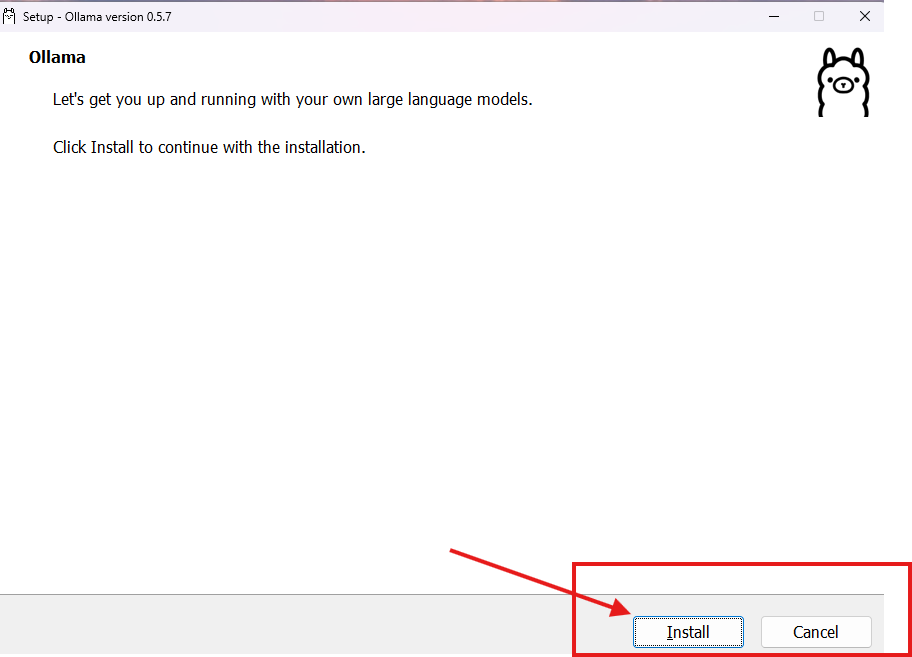


Github：



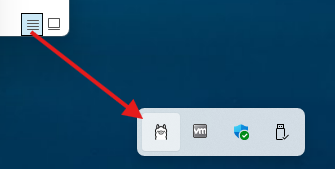
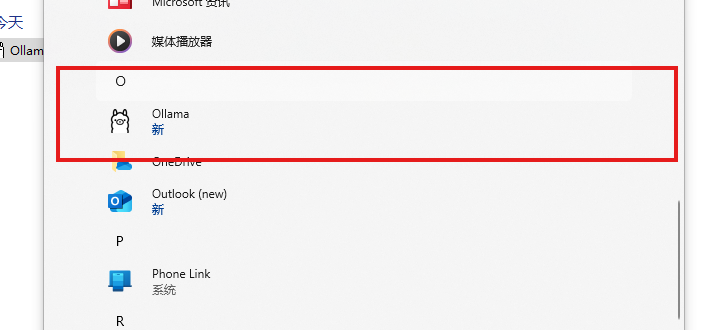
2：点击下载完之后会找到下载文件的地址点击安装





3：安装完之后会在开始的界面中出现该软件的图标：

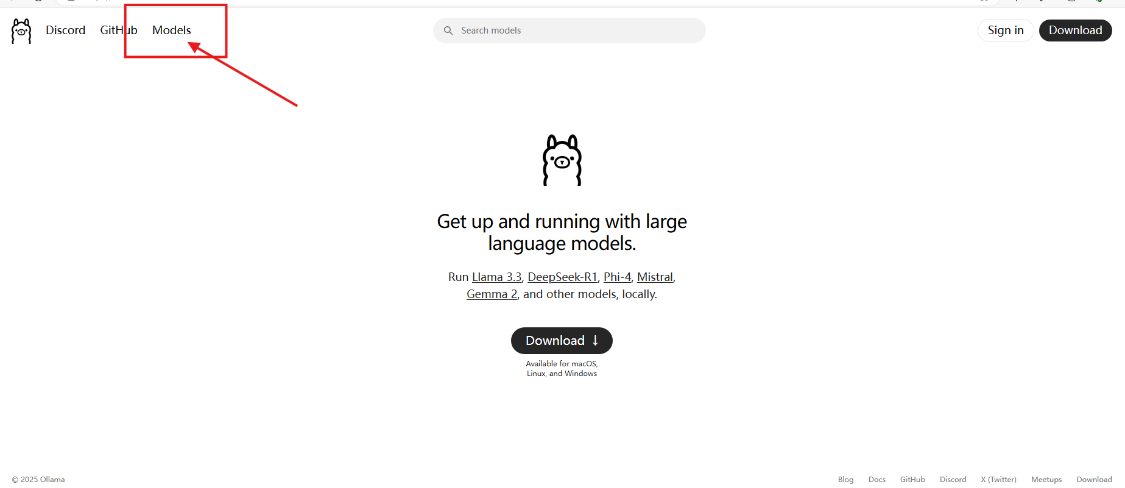
（由于目前应该只支持直接在c盘中安装，所以本文就直接安装在c盘中，如果可以安装在其他盘请@我）



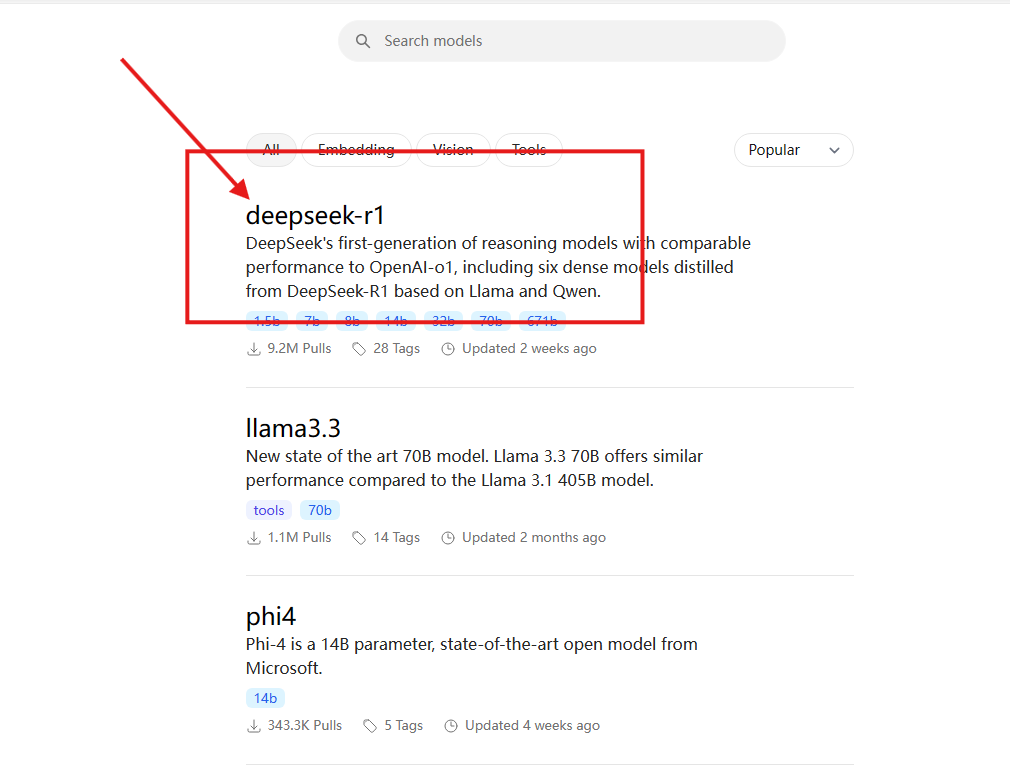
出现这个图标就说明安装成功了

4：我们回到官网，点击Models选择我们要安装的版本：

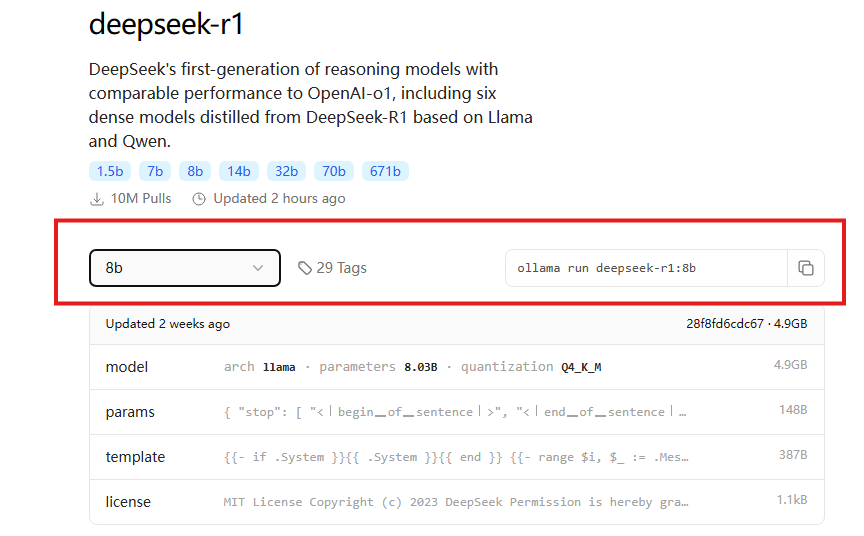
1）



2）

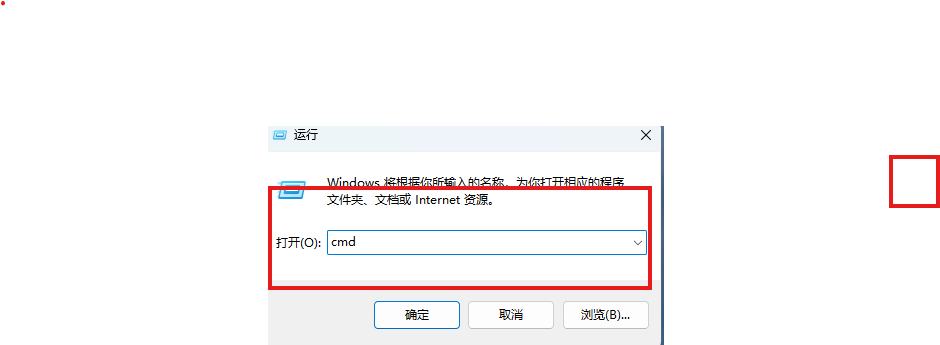


3）本文安装的14b的版本，大小为9GB（需要在c盘中预留出对应版本所需要的空间）

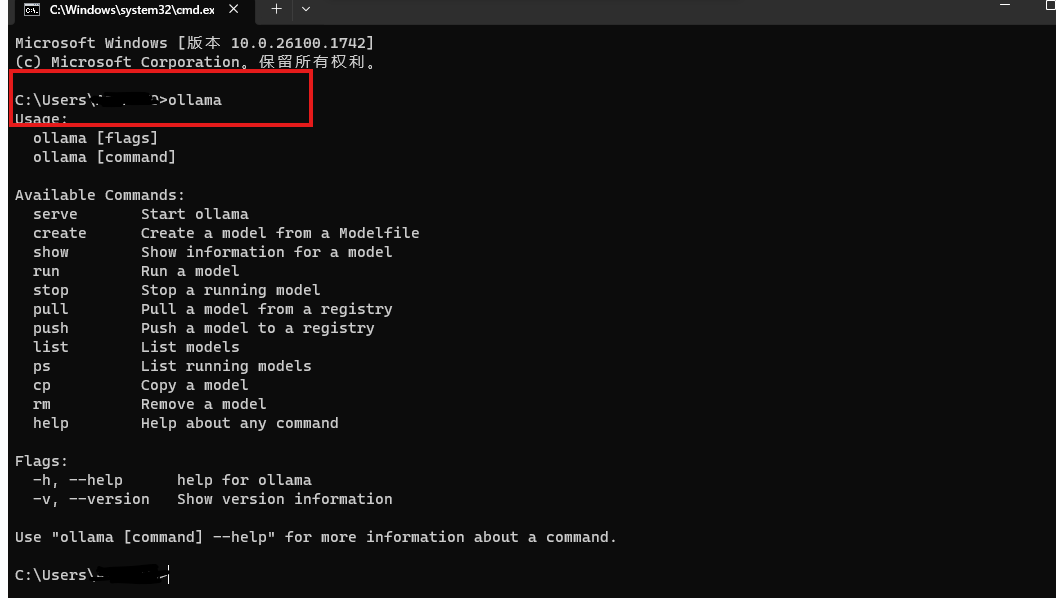


5：选定好版本后我们使用win+r打开cmd，输入命令“ollama”

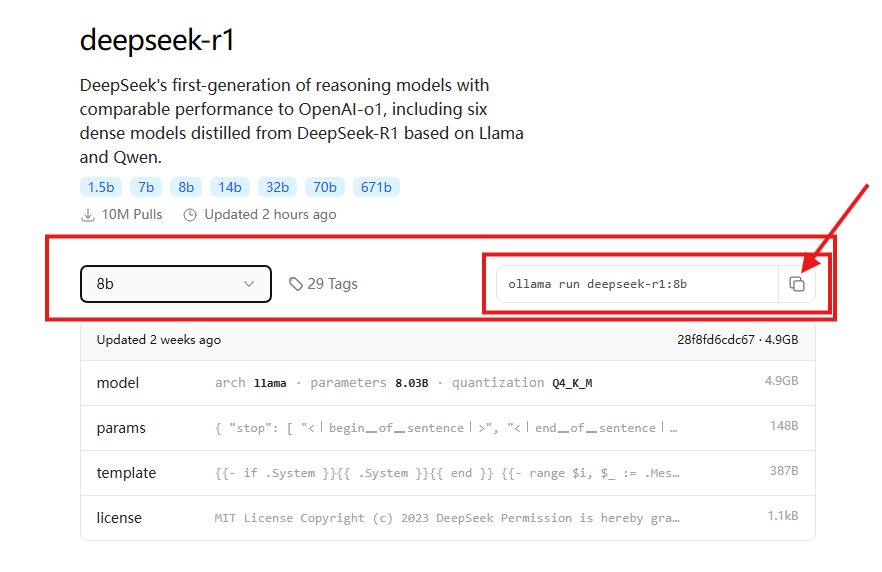
1）

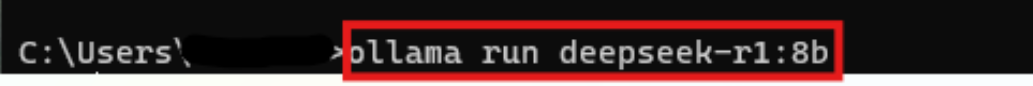


2）

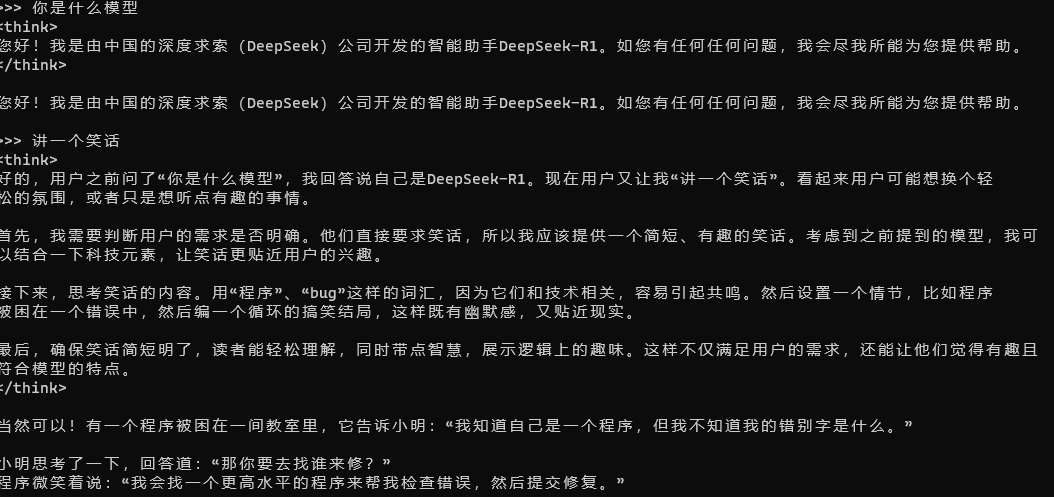


6：输入完命令之后回到官网复制对应版本下载所需要的命令，将命令在cmd中执行





7：下载完之后就可以跟deepseek进行简单的对话了



## 附加

**但是cmd中跟他对话太过抽象所以我们需要一个简单的界面完成我们的对话**

**1：安装Docker**

**官网：**[Docker: Accelerated Container Application Development](https://www.docker.com/)

**介绍docker：**

**\*\*Docker\*\* 是一种开源的容器化平台，允许你将软件包装成可运行的镜像，并通过容器来轻量级地运行这些应用程序。**

**它的核心思想是“将软件打包成可重复使用的镜像”，这样可以简化开发、部署和管理过程，同时确保环境的一致性。**

**### 主要特点：**

**1. \*\*轻量级\*\*：Docker 容器比传统虚拟机更轻量，启动速度快。**

**2. \*\*可重复性\*\*：通过镜像，可以在任何环境下重现相同的配置和设置。**

**3. \*\*跨平台支持\*\*：Docker 容器可以运行在 Linux、Windows、macOS 等多种操作系统上。**

**4. \*\*依赖管理\*\*：自动处理依赖关系，减少手动配置。**

**### 核心概念：**

**- \*\*镜像（Image）\*\*：包含应用程序和其所有依赖的可执行文件包装。**

**- \*\*容器（Container）\*\*：运行镜像创建的轻量级进程，可以共享操作系统层面的资源。**

**- \*\*构建工具（Dockerfile）\*\*：定义镜像的内容，包括软件、编译器、依赖库等。**

**- \*\*运行命令\*\*：**

**- `docker build`：从 Dockerfile 构建镜像。**

**- `docker run`：运行镜像创建的容器。**

**- `docker stop`：停止并删除容器。**

**- `docker commit`：将容器中的更改保存为新的镜像。**

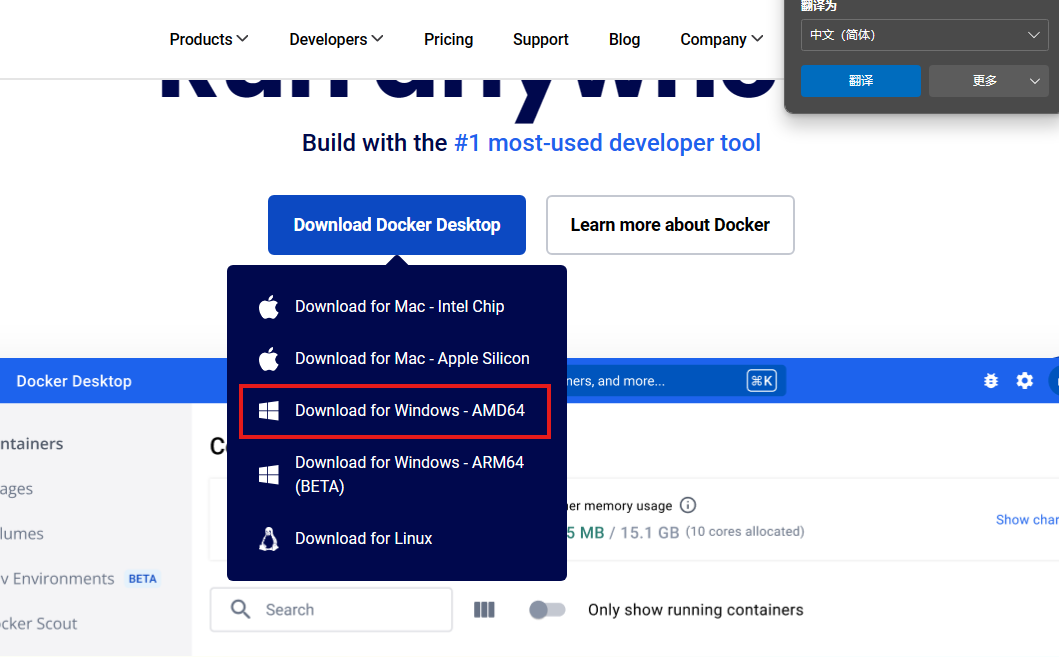
**### 常见用途：**

**- \*\*开发与测试\*\*：简化环境一致性，快速重建和运行应用。**

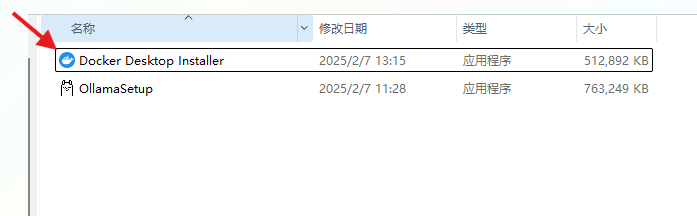
**- \*\*部署\*\*：通过容器轻松部署应用程序，无需物理服务器。**

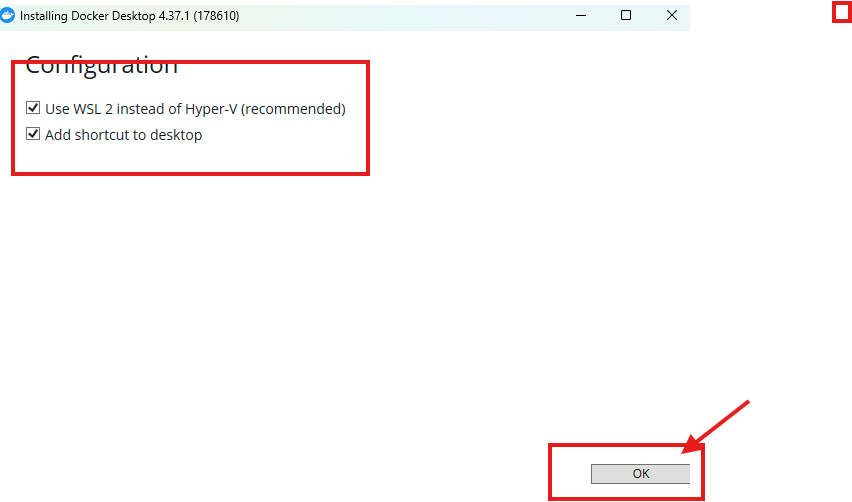
**- \*\*持续集成/持续交付（CI/CD）\*\*：自动化构建、测试和部署流程。**

1）进入官网之后选择windows版本的进行下载

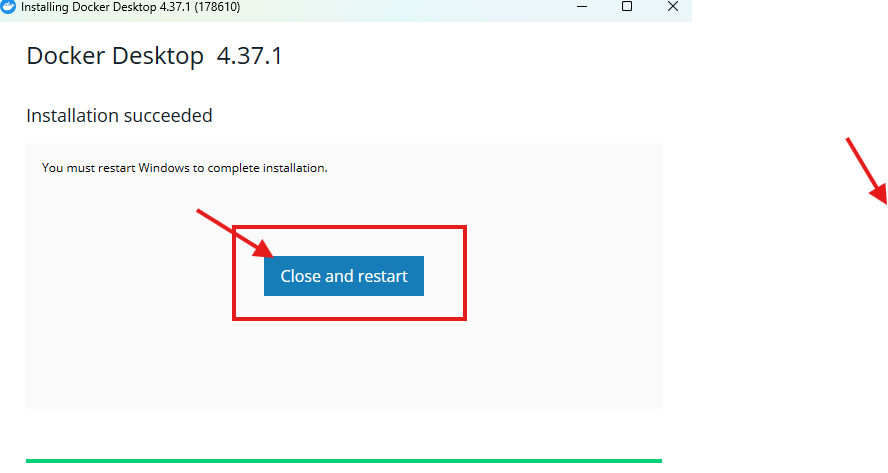


2）下载完之后进行安装

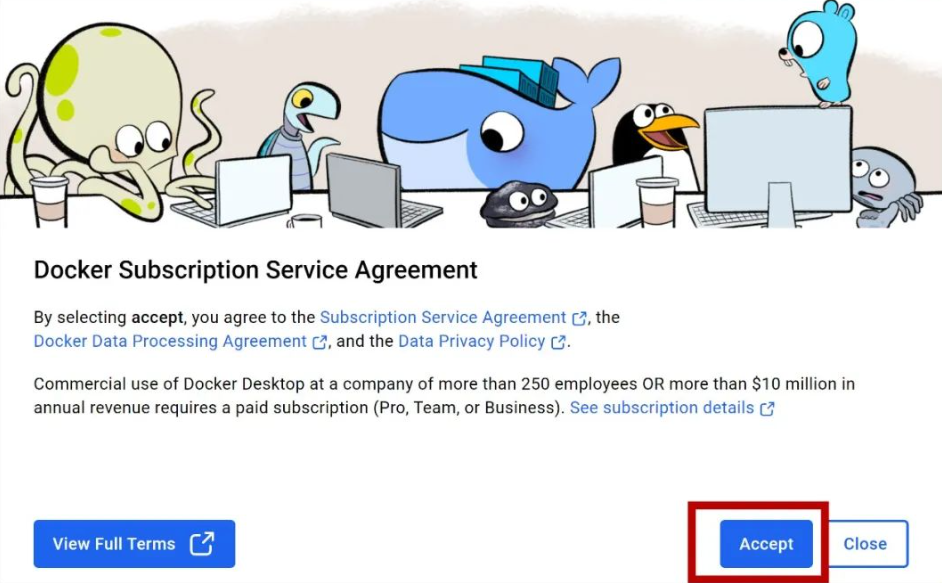




**\*\*点击完之后电脑会重新启动\*\***



重启后我们接受条款就可以使用Docker



点击完接受之后会出选则：

原文：Use recommended settings (requires administrator password)

Docker Desktop automatically sets the necessary configurations that work for most developers.

含义：使用建议的设置（需要管理员密码）

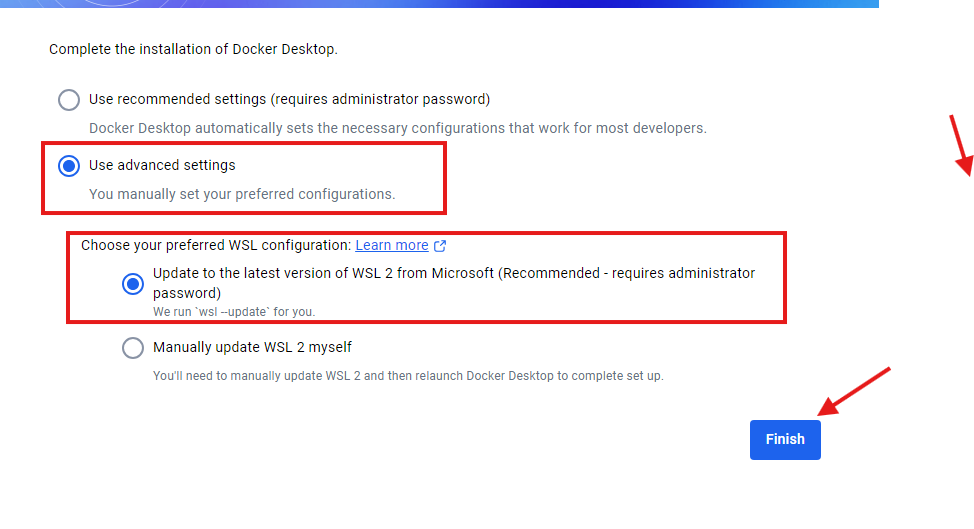
Docker Desktop 会自动设置适用于大多数开发人员的必要配置。

原文：Use advanced settings

You manually set your preferred configurations.

含义：使用高级设置

您可以手动设置首选配置。

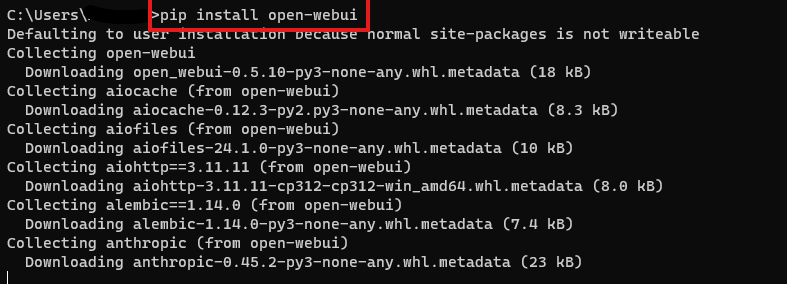


**（由于我使用的虚拟机安装的本程序他在虚拟机里面开不了虚拟化支持所以Docker在打开时会进行报错，如果大家想继续应用Docker可以参考：**<https://blog.csdn.net/2401_84204207/article/details/145389919>**，大家可以接着往下看下面会有别的方法在简洁的页面中打开deepseek）**

**2：安装Open WebUI**

**在终端或命令行工具中运行以下命令，安装 Open-WebUI：“pip install open-webui”**

**如果 pip 版本较低，可以更新下“python -m pip install --upgrade pip”**



**（如果大家“pip install open-webui”命令可以直接执行完毕，大家可以参考：**<https://blog.csdn.net/m0_65555479/article/details/145389847>，进行之后的可视化部署）

**3：安装ChatBox**

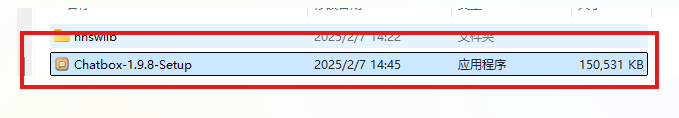
**官网：<https://chatboxai.app/zh>**

**1）进入网站之后点击下载之后他会自动开始下载**

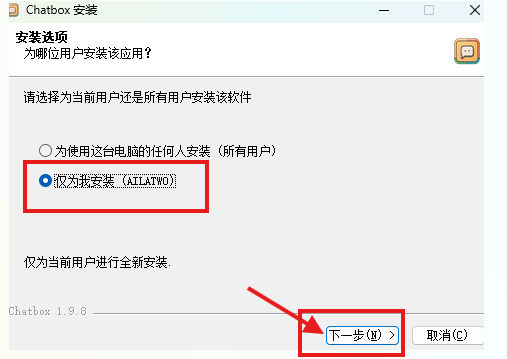




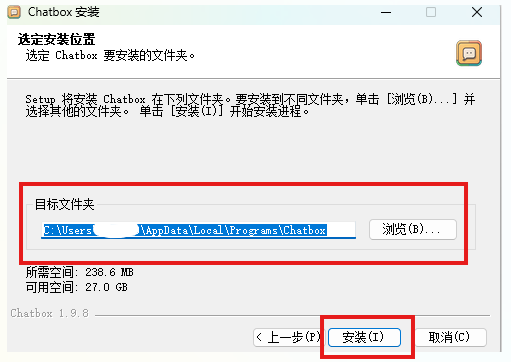
**2）打开安装程序**



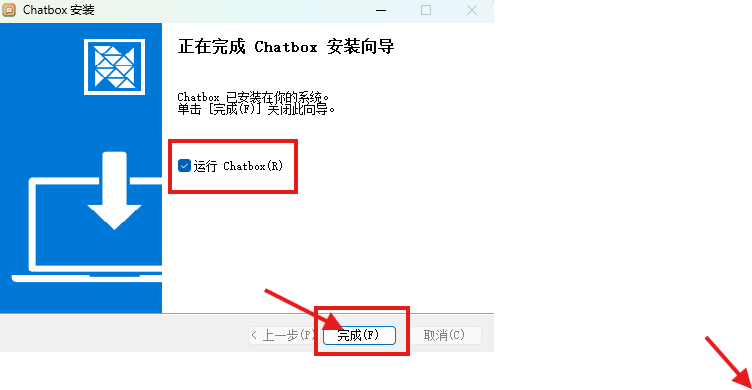
**3）选择仅为我安装**



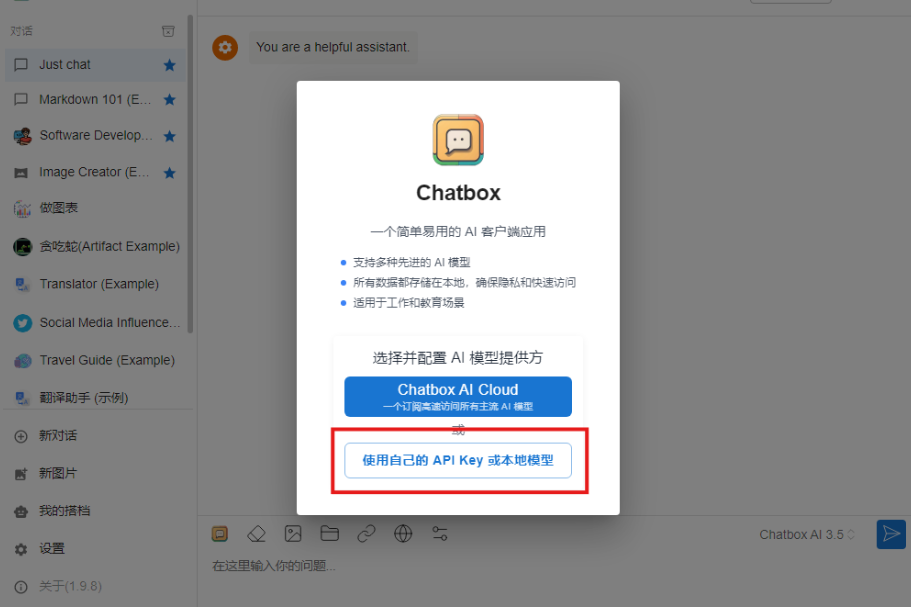
**4）选择好安装的地址点击安装**



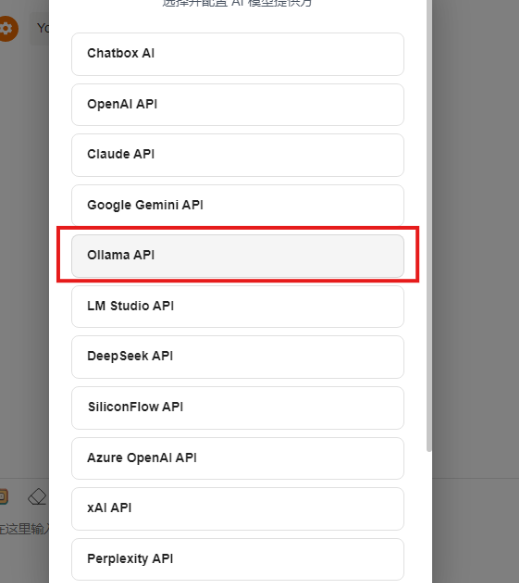
**5）安装完之后点击完后它会自动运行**



**6）运行程序之后我们可以选择运行自己本地模型：**

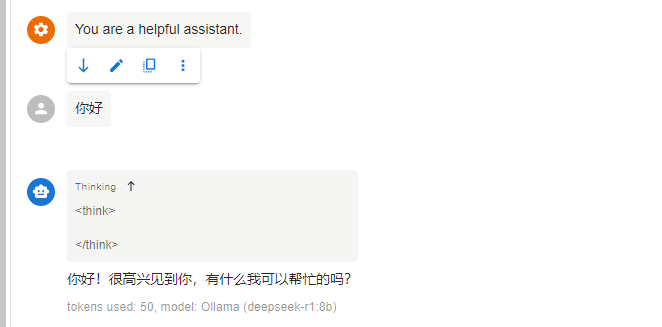


**7）我们选择Ollama API**



**8）进入之后我们选择我们在本地部署的模型版本，之后点击保存就可以进行对话了**





参考文献：

[高效快速教你deepseek如何进行本地部署并且可视化对话-腾讯云开发者社区-腾讯云](https://cloud.tencent.com/developer/article/2492444)

<https://blog.csdn.net/2401_84204207/article/details/145389919>

<https://blog.csdn.net/m0_65555479/article/details/145389847>