

CDM LAB 环境配置

本课程的所有LAB都需要使用Linux环境，使用Windows或MacOS的同学需要安装虚拟机

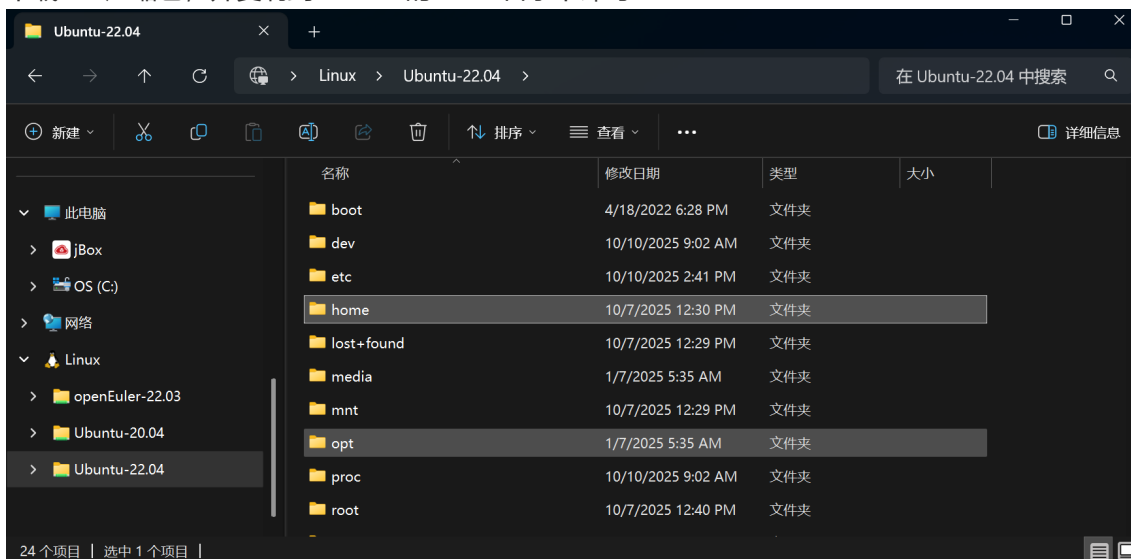
Windows

方案一：使用WSL（推荐，更轻量更快速）

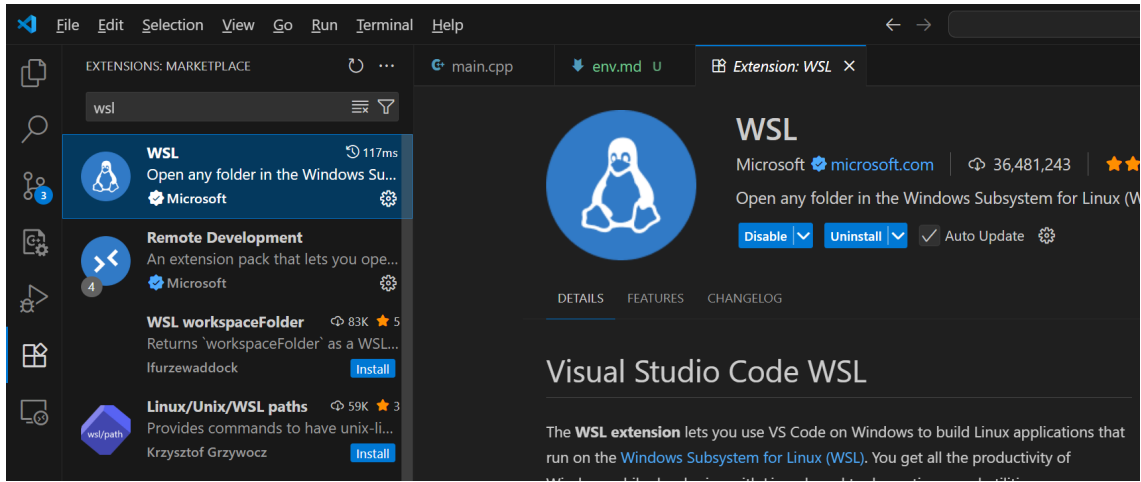
1. 以管理员身份打开 PowerShell 或 Windows Terminal。
2. 执行以下命令：

```
wsl --install -d Ubuntu-22.04
```

3. 系统将自动：
 - a. 启用 WSL 和虚拟机平台功能。
 - b. 下载并安装最新 WSL2 内核。
 - c. 安装对应的Linux系统（为便于后续环境配置，请确认安装的是Ubuntu22.04）。过程中需要设置用户名和密码。
4. 安装结束后会自动进入Ubuntu终端，使用exit命令退出。以后每次使用只需要在PowerShell 或 Windows Terminal中键入wsl命令即可启动Ubuntu系统。
5. WSL的文件可以直接在Windows文件资源管理器中访问。每次发布lab时，使用Windows在Canvas下载LAB压缩包，并复制到Ubuntu的home目录下即可

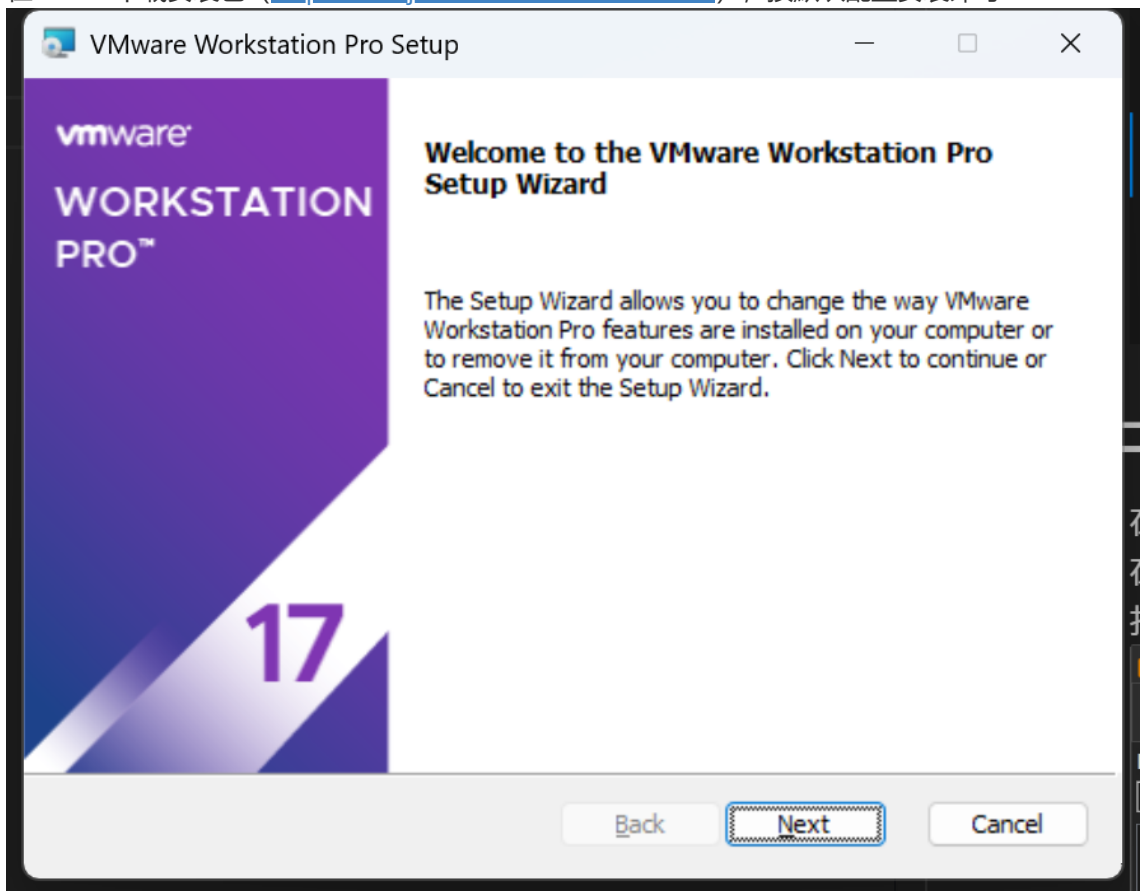


6. (可选) 安装vscode, 并安装WSL扩展, 随后连接到WSL以便于开发。



方案二：使用VMware Workstation（更方便）

1. 在canvas下载安装包 (<https://oc.sjtu.edu.cn/files/11744800>) , 按默认配置安装即可

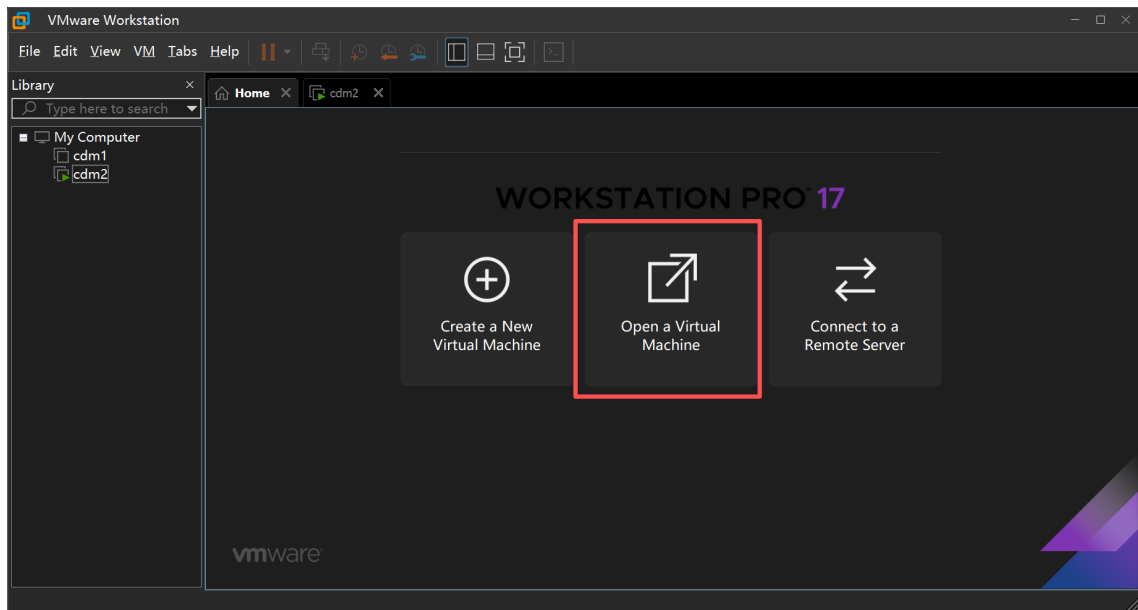


2. 下载虚拟机镜像 UbuntuCDM.ova

链接: <https://pan.sjtu.edu.cn/web/share/bfdf699da9a9a0d0ff2d4838bb0e80ff>

提取码: 3w9z

3. 打开软件，选择“打开虚拟机”，选择刚才下载的虚拟机镜像，等待导入



4. 完成后你就得到了一台Ubuntu22.04的虚拟机，默认用户名密码都是“cdm”

MacOS

请参考目录下的 env mac-arm.pptx

附录

A: VMware自带软件介绍

• 软件使用



– 浏览器：Ubuntu自带了火狐浏览器，你可以通过浏览器登录 canvas 下载和上传作业



– 代码编辑器：我们为你预装了VS Code编辑器，你可以使用这个打开代码文件夹

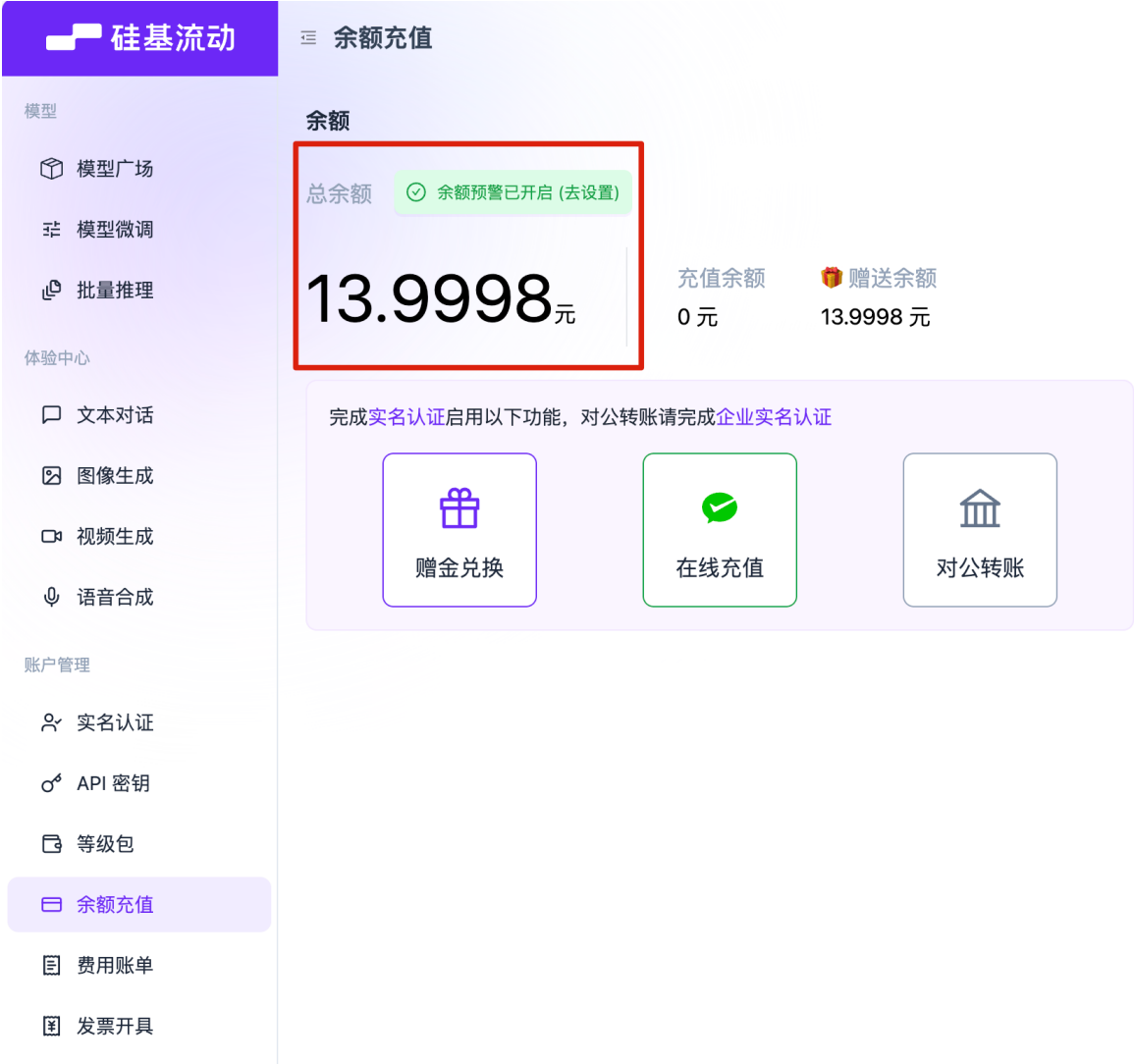


– 应用商店：你可以使用Ubuntu自带的应用商店下载应用或者也可以通过浏览器搜索下载自己需要的应用

B: 如何在代码中通过 API 使用 LLM

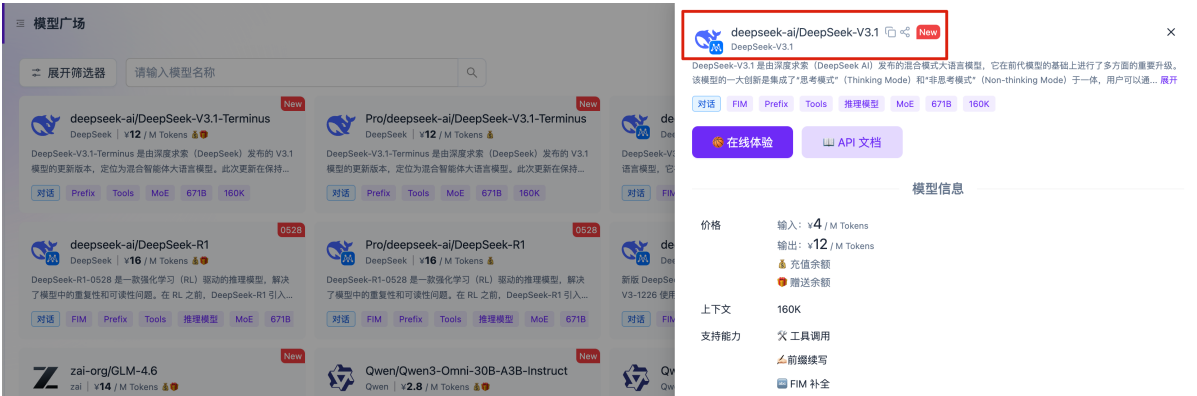
同学们可以自由使用自己知道的 LLM 服务，但如果你还不知道在哪里可以访问这样的服务，可以按照本教程进行操作：

硅基流动是一家国内的 LLM 服务分发平台，以 API 的形式提供许多开源模型的访问。它为每位新用户提供了约 2000 万 tokens 的免费额度，足够本课程的 LAB 使用。同学们在 <https://cloud.siliconflow.cn/> 注册登录后可以前往余额见面查看剩余额度



登陆硅基流动平台后，可以在个人中心 -> API 密钥这里创建并复制自己的密钥，进行 LAB 时需要填写到 `llm.cpp` 的 `kApiKey` 常量中

在[模型广场](#)可以看到硅基流动提供访问的模型，选择完毕后复制其名字用于在 LAB 框架代码中指定模型



此外，由于免费用户在 LLM Serving 时调度优先级较低，同学们可能经常遇到服务忙的问题，此时可以在几家服务商之间切换，这里为大家罗列一些：

- [阿里云百炼平台](#) 新用户获赠一些 Qwen 模型免费额度
- [豆包方舟](#) 新用户获赠部分 Doubao 模型免费额度
- [Google AI Studio](#) 可以免费使用部分 Gemini 模型

Note: 切换平台时需要同时修改 `kApiKey` 和 `kApiBase` 常量

C: VSCode中使用大模型辅助编程 (猜你喜欢)

可以在 VSCode->Extensions 中安装 Github Copilot, 免费用户拥有一定额度的代码补全和 AI 问答次数。