

部署命令

Stephen CUI

2023-09-12

1 下载权重

```
sudo apt-get update
```

安装 git-lfs 因为较大文件 git 不管理，需要使用 git-lfs 下载，但是下载较慢，建议在[清华云盘](#)中下载。

```
sudo apt-get install git-lfs
```

2 安装 Python 与环境创建

这个命令可以安装 Python 其他版本，系统带有 Python3.8，可以不安装

```
sudo apt-get install python3.10
```

% 删除默认的软连接，因为默认 python 链接到 python2

```
sudo rm -rf /usr/bin/python3
```

% 将python 默认设置为 Python3.8, Python3.8 应该是 Ubuntu20 默认安装的版本

```
sudo ln -s /usr/bin/python3.8 /usr/bin/python
```

安装 Python3 的 pip

```
sudo apt-get install python3-pip
```

可以不运行，除非版本不够（升级至最新的 pip）：

```
pip install --upgrade pip
```

从 Python 3.6 开始，创建虚拟环境的推荐方法是使用 venv 模块。因此我们先要安装提供 venv 模块的 python3-venv 软件包。运行命令：

```
sudo apt install python3-venv
```

切换项目目录。在目录中，运行

```
python -m venv llm
```

命令来创建新的虚拟环境，llm 可以更改为任意环境名。

要开始使用此虚拟环境，您需要通过运行 `activate` 脚本将其激活。`source` 命令将会加载 `python` 的虚拟环境

```
source llm/bin/activate
% conda activate llm
```

一旦激活，虚拟环境的 `bin` 目录将添加到 `PATH` 变量的开头。此外，您的 `Shell` 提示符也会更改，并且会显示您当前正在使用的虚拟环境的名称。

使用 `Python` 包管理器 `pip` 安装 `python` 包，在虚拟环境中，可以使用命令 `pip` 代替 `pip3`，并使用 `python` 代替 `python3`

```
pip install numpy -i https://pypi.tuna.tsinghua.edu.cn/simple
```

完成工作后停用虚拟环境，只需键入 `deactivate`，您将返回到常规 `shell`。

```
deactivate
```

必须在 `Git Bash` 中运行

```
GIT_LFS_SKIP_SMUDGE=1 git clone https://huggingface.co/THUDM/chatglm-6b
```

3 网页版：基于 `gradio` 和 `streamlit`

`Gradio` 是通过友好的 `Web` 界面演示机器学习模型的最快方式，以便任何人都可以在任何地方使用它！

```
pip install gradio mdtex2html
```

`Gradio` 直接使用 `Python` 运行，国内访问速度慢。

```
python web_deom.py
```

建议使用 `streamlit`，安装命令：

```
pip install streamlit
pip install streamlit-chat
```

```
streamlit run web_demo2.py
```

4 微调模型

4.1 软件依赖

```
pip install rouge_chinese nltk jieba datasets
```

`cuda` 更多版本

更新 `cuda` 使用以下命令

```
wget https://developer.download.nvidia.com/compute/cuda/11.7.1/local_installers/cuda_11.7.1_515
```

```
sudo sh cuda_11.7.0_515.43.04_linux.run
```

如果你安装了驱动，在安装的时候需要取消驱动的勾选：

```
apt install ubuntu-drivers-common
```

```
ubuntu-drivers devices
```

执行微调的文件，用以下命令：

```
bash train.sh
```

使用 SFTP 传输本地文件的配置：remotePath, host, password, name 按实际修改，ignore 为不传输文件

```
{
  "name": "ECS-aeqk",
  "host": "101.126.37.159",
  "protocol": "sftp",
  "port": 22,
  "username": "root",
  "password": "Watone_passwd123!@#",
  "remotePath": "llm/glm",
  "uploadOnSave": true,
  "useTempFile": false,
  "openSsh": false,
  "ignore": [
    "**/.vscode/**",
    "**/.git/**",
    "**/.DS_Store/**",
    "**/__pycache__/**"
  ]
}
```

5 参考链接

[如何在 Ubuntu 18.04 创建 Python 虚拟环境](#)

[yizhongw/self-instruct](#)

[清华大学GLM-6B](#)

[模型权重所在云盘位置](#)