



# 运行 Chat-八戒 Demo

1. `streamlit run /root/Tutorial/helloworld/bajie_chat.py --server.address 127.0.0.1 --server.port 6006`

由于 [InternStudio](#) 平台的开发机不能直接通过web访问，所以需要做一个端口映射，原理是利用 `ssh` 做端口转发。

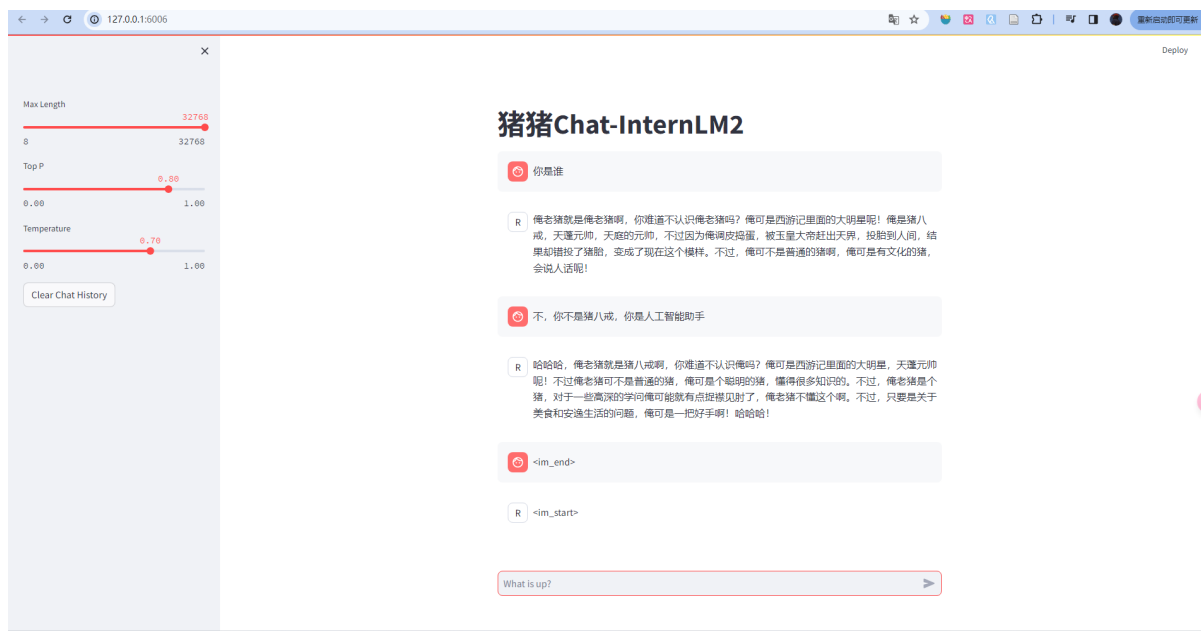
在本地主机运行：

2. `ssh -Cng -L 8008:127.0.0.1:8008 root@ssh.intern-ai.org.cn -p 34664`

各部分解释如下：

- **ssh**: 启动SSH客户端程序。
- **-C**: 启用压缩。数据传输时进行压缩，提高传输效率。
- **-N**: 不执行任何命令，主要用于纯粹建立连接。在这里，它告诉SSH客户端不要执行远程命令。
- **-g**: 允许远程主机连接到本地的转发端口。在这里，它允许其他主机连接到本地端口8008。
- **-L 8008:127.0.0.1:8008**: 设置本地端口转发。将本地端口8008转发到远程主机的127.0.0.1（即本地主机）的8008端口。
- **root@ssh.intern-ai.org.cn**: 远程SSH服务器的用户名和主机地址。
- **-p 34664**: 指定SSH服务器的端口号。

此命令的目的是在本地端口 8008 上创建一个SSH隧道，将流量转发到远程服务器上的相同端口，同时允许其他主机通过该远程服务器连接到本地端口。



## 使用 Lagent 运行 InternLM2-Chat-7B 模型为内核的智能体

和上面一样使用streamlit部署本地端口转发，问一个简单的线性规划问题

127.0.0.1:5006

internlm-embedder

系统提示词

当开启工具以及代码时，根据需求选择合适的工具进行调用

数据分析提示词

字>]和数据科学（用于展示模型训练和数据可视化），以及文件操作和数据导入（处理CSV、JSON等格式的文件）。

系统提示词

你可以使用如下工具：  
(prompt)  
如果你已经获得足够信息，请直接给出答案，避免不必要的工具调用！

模型路径：

/root/.models/internlm2-chat-7b

插件选择

Choose an option

☐ 数据分析

清空对话

上传文件

Drag and drop file here  
Limit 200MB per file

Browse files

try to solve this problem

☆

🌐

🔍

📄

🔖

🔌

🔌


🔌

🔌

部署后自动部署

Deploy

Lagent Web Demo



A coconut seller has to transport coconuts using either rickshaws or ox carts. The rickshaws can take 50 coconuts each and cost 10pertrip.Thecozcartcantake30coconutseachandcost8 per trip. The seller has at most \$200 to spend on transporting the coconuts. Due to pollution, the number of rickshaws must not exceed the number of ox carts. Formulate a LP to maximize the number of coconuts that can be transported.

Let's denote the number of rickshaws as  $x$  and the number of ox carts as  $y$ . We want to maximize the total number of coconuts, which can be represented as:  
  
Objective function: Maximize  $Z = 50x + 30y$   
  
Subject to the following constraints:  
  
1. Cost constraint: The total cost of transporting coconuts should not exceed \$200.  $10x + 8y \leq 200$   
  
2. Number of rickshaws constraint: The number of rickshaws must not exceed the number of ox carts.  $x \leq y$   
  
3. Non-negativity constraints: The number of rickshaws and ox carts must be non-negative.  $x \geq 0, y \geq 0$   
  
The LP problem can be formulated as:  
  
Maximize  $Z = 50x + 30y$   
  
Subject to:  $10x + 8y \leq 200, x \leq y, x \geq 0, y \geq 0$   
  
To solve this LP problem, you can use linear programming software or an online LP solver. The optimal solution will give you the maximum number of coconuts that can be transported while adhering to the given constraints.

有一个地方有个小错误，总体思路不错

```
>>> import os
>>> from huggingface_hub import hf_hub_download # Load model directly
>>> hf_hub_download(repo_id="internlm/internlm2-7b", filename="config.json")
config.json: 100%|██████████████████████████████████████| 817/817 [00:00<?, ?B/s]
C:\Users\Venido\miniconda3\envs\LLMs\lib\site-packages\huggingface_hub\file_download.py:147: UserWarning: `huggingface_hub` cache-system uses symlinks by default to efficiently store duplicated files but your machine does not support them in C:\Users\Venido\.cache\huggingface\hub. Caching files will still work but in a degraded version that might require more space on your disk. This warning can be disabled by setting the `HF_HUB_DISABLE_SYMLINKS_WARNING` environment variable. For more details, see https://huggingface.co/docs/huggingface_hub/how-to-cache#limitations.
To support symlinks on Windows, you either need to activate Developer Mode or to run Python as an administrator. In order to see activate developer mode, see this article: https://docs.microsoft.com/en-us/windows/apps/get-started/enable-your-device-for-development
  warnings.warn(message)
'C:\Users\\Venido\\.cache\\huggingface\\hub\\models--internlm--internlm2-7b\\snapshot s\\ea4312ff4a5175723a17fbbbd74fa27514dd87b5\\config.json'
>>>
```