

```

(base) root@intern-studio:~# conda env list
# conda environments:
#
base                  * /root/.conda
internlm-demo         /root/.conda/envs/internlm-demo
xcomposer-demo        /root/.conda/envs/xcomposer-demo

(base) root@intern-studio:~# conda activate internlm-demo
(internlm-demo) root@intern-studio:~# python /root/code/InternLM/cli_demo.py
Loading checkpoint shards: 100%|
=====Welcome to InternLM chatbot, type 'exit' to exit=====
User >>> 请写一段<<300>>>字的故事，故事的梗概如下：<<<“在一个未来科技高度发达的社会中，人们生活在智能城市中，生活方便且高度自动化。然而，一天，一位年轻的工程师发现了一个被系统隐藏的密码。这个密码是什么？它又是如何影响城市人们的生活？年轻工程师将如何发现这个发现？”>>>，请开始编写。
robot >>> 在一个未来科技高度发达的社会中，人们生活在智能城市中，生活方便且高度自动化。他们通过智能手机、智能家居等科技设备完成日常生活所需的一切，甚至可以通过虚拟现实技术进行远程办公和学习。但是，一切都平静地持续着直到有一天，一位年轻的工程师在一家大型科技公司的开发中心发现了一个被系统隐藏的密码。

据悉，这家公司曾经研发了一种名为“智能梦境”的高级软件，该软件可以通过监控用户的梦境来实现对用户的心理和行为的分析，并将这些数据传输给第三方机构进行分析。然而，随着软件的不断升级和完善，该软件开始出现了一个被系统隐藏的功能——“梦境共享”。

这意味着，用户的梦境不仅仅是可以被监控和分析，还可以被他人共享。这个功能的出现引发了广泛的恐慌和不安，许多人都开始担心自己的隐私权会被侵犯。年轻的工程师决定调查这个秘密，并且发现这个秘密已经影响到了社会和人们的生活。

为了保护用户的隐私权，年轻的工程师决定向公众揭露这个秘密，并且向相关部门提出了一系列的建议和呼吁。然而，这个决定引发了一系列的问题和矛盾。一方面，许多人开始反对这个决定，他们担心自己的隐私权会被泄露，甚至有人开始攻击和恐吓年轻的工程师。另一方面，公司的高层却认为这个决定会对公司的利益产生负面影响，因此拒绝公开这个秘密。

在经历了各种困难和挑战之后，年轻的工程师终于找到了一合适的时机和平台来公开这个秘密。他通过社交媒体和新闻媒体向公众揭露了这个秘密，并且呼吁社会各界共同关注和保护用户的隐私权。最终，这个秘密被揭开了。科技公司和政府也采取了了一系列措施来保护用户的隐私权。智能城市的未来也因此显得更加光明和美好。

```

2. Huggingface下载 (失败)

```

1 # huggingface 下载方法，但是InternLM用这种方法下载似乎会失败：HTTPError: Response
  details: {'Code': 10010205001, 'Message': '获取模型版本失败，信息: record not
  found', 'RequestId': '3a020967-3af6-401c-9a24-9cc8b73cb7e5', 'Success':
  False}, Request id: d410724cfe7542d5930d3aaa4d4eb803
2
3 import torch
4 from transformers import AutoTokenizer, AutoModelForCausalLM
5 tokenizer = AutoTokenizer.from_pretrained("internlm/internlm-chat-20b",
  cache_dir='/root/model', trust_remote_code=True)
6 # Set `torch_dtype=torch.bfloat16` to load model in bfloat16, otherwise it
  will be loaded as float32 and cause OOM Error.
7 model = AutoModelForCausalLM.from_pretrained("internlm/internlm-chat-20b",
  torch_dtype=torch.bfloat16, trust_remote_code=True).cuda()
8 model = model.eval()
9 output, history = model.chat(tokenizer, "Hello! Today is sunny, it is time
  to go out")
10 print(output)
11 # Hello! Today is sunny, and it sounds like a great day to go out an enjoy
  the weather. What would you like to do?

```

3. huggingface_hub下载

```

1 # 可以继续未完成的下载过程，并保存到本地目录/content/model
2 huggingface-cli download --resume-download internlm/internlm-20b --local-
  dir /content/model

```

4. openxlab镜像下载：

```

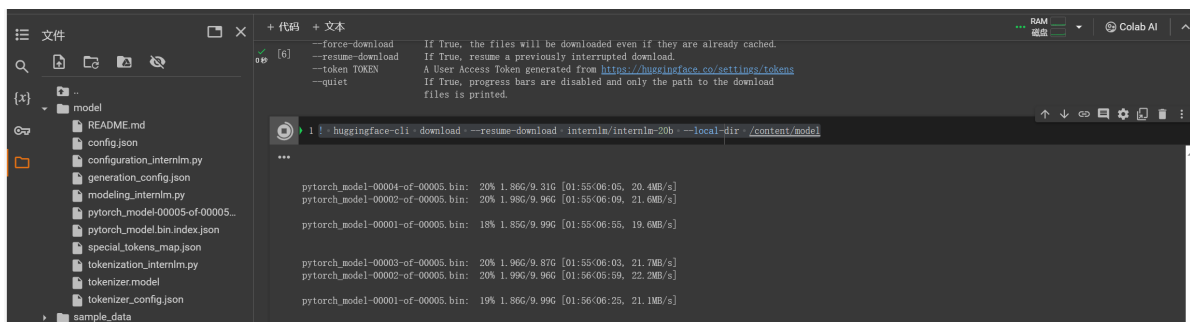
1 from openxlab.model import download
2 download(model_repo='OpenLM/InternLM-20b', model_name='InternLM-20b',
  output='/content/model', overwrite='No')

```

5. 手动下载：在抱抱脸上直接找到模型权重文件，摁下。

2.3 下载结果

huggingface_hub下载 (config.json)：



2.4 关于huggingface下载的知识延伸

1. ModelScope下载 (snapshot_download) :

官方链接: [HfApi.Client\(huggingface.co\)](https://huggingface.co/HfApi.Client)

1. 通常是国内源下载。
2. 比较了令人困惑的是`repo_id`, 可以在这里找到:

书生·浦语-对话-20B [Shanghai_AI_Laboratory/internlm-chat-20b](#)



模型介绍



模型文件



提交反馈

Notebook快速开发

3. `cache_dir` : 模型保存位置, 建议设置。

2. huggingface-cli下载:

这是一个多层的参数结构:

先看看 `huggingface-cli --help`:

```
1  usage: huggingface-cli <command> [<args>]
2
3  positional arguments:
4    {env,login,whoami,logout,repo,upload,download,lfs-enable-largefiles,lfs-
      multipart-upload,scan-cache,delete-cache}
5
6    huggingface-cli command helpers
7    env                Print information about the environment.
8    login              Log in using a token from
      huggingface.co/settings/tokens
9    whoami             Find out which huggingface.co account you are logged
      in as.
10   logout             Log out
11   repo               {create} Commands to interact with your
      huggingface.co repos.
12   upload             Upload a file or a folder to a repo on the Hub
13   download           Download files from the Hub
14   lfs-enable-largefiles
      Configure your repository to enable upload of files
      > 5GB.
15   scan-cache         Scan cache directory.
16   delete-cache       Delete revisions from the cache directory.
17
18  options:
19    -h, --help         show this help message and exit
20
```

再看看 `huggingface download --help`。这里的**positional arguments**不需要带关键字, 且**顺序不可变**。其中`repo_id`是huggingface的model中的仓库路径, 如`internlm/internlm-20b`:

[internlm/](#) **internlm-20b**
[like](#) 67

[Text Generation](#)

[Transformers](#)

[PyTorch](#)
[internlm](#)
[feature-extraction](#)
[custom_code](#)

[License: apache-2.0](#)

```

1  usage: huggingface-cli <command> [<args>] download [-h] [--repo-type
2                                     {model,dataset,space}]
3                                     [--revision REVISION] [--
4                                     include [INCLUDE ...]]
5                                     [--exclude [EXCLUDE ...]]
6                                     [--cache-dir CACHE_DIR]
7                                     [--local-dir LOCAL_DIR]
8                                     [--local-dir-use-symlinks
9                                     {auto,True,False}]
10                                    [--force-download] [--
11                                    resume-download]
12                                    [--token TOKEN] [--quiet]
13                                    repo_id [filenames ...]
14
15 positional arguments:
16   repo_id                ID of the repo to download from (e.g.
17   `username/repo-name`).
18   filenames              Files to download (e.g. `config.json`,
19   `data/metadata.jsonl`).
20
21 options:
22   -h, --help            show this help message and exit
23   --repo-type {model,dataset,space}
24                           Type of repo to download from (e.g. `dataset`).
25   --revision REVISION    An optional Git revision id which can be a branch
26   name, a tag, or a commit hash.
27   --include [INCLUDE ...]
28                           Glob patterns to match files to download.
29   --exclude [EXCLUDE ...]
30                           Glob patterns to exclude from files to download.
31   --cache-dir CACHE_DIR
32                           Path to the directory where to save the downloaded
33   files.
34   --local-dir LOCAL_DIR
35                           If set, the downloaded file will be placed under
36   this directory either as
37                           a symlink (default) or a regular file. Check out
38   https://huggingface.co/docs/huggingface_hub/guides/download#download-
39   files-to-local-folder for more details.
40   --local-dir-use-symlinks {auto,True,False}
41                           To be used with `local_dir`. If set to 'auto', the
42   cache directory will be
43   used and the file will be either duplicated or
44   symlinked to the local
45   directory depending on its size. If set to `True`, a
46   symlink will be
47   created, no matter the file size. If set to `False`,
48   the file will either

```

```
36 | be duplicated from cache (if already exists) or
    | downloaded from the Hub
37 | and not cached.
38 | --force-download    If True, the files will be downloaded even if they
    | are already cached.
39 | --resume-download   If True, resume a previously interrupted download.
40 | --token TOKEN       A User Access Token generated from
    | https://huggingface.co/settings/tokens
41 | --quiet             If True, progress bars are disabled and only the
    | path to the download
42 | files is printed.
```

3. openxlab镜像下载:

文档: [Download Models | Docs-OpenXLab](#)

无他,就是要配置一下AK和SK,根据[教程](#)走一遍就好。当然这个配置过程在google colab上可能有些繁琐,不是很建议。其他的服务器可以一试。

ps: 这里挖个坑

3 完成浦语·灵笔的图文理解及创作部署 (需截图)

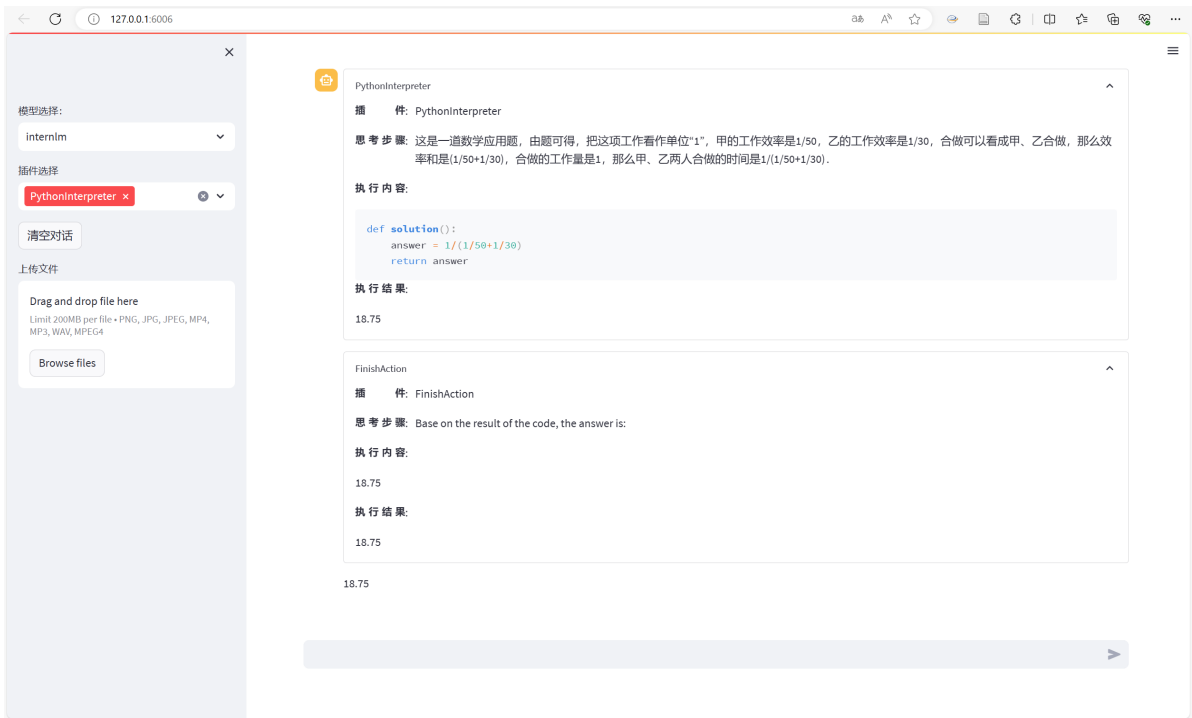
3.1 运行命令 (环境配置略, 直接参考[链接](#)):

```
1 | conda activate internlm-demo
2 | streamlit run /root/code/lagent/examples/react_web_demo.py --server.address
   | 127.0.0.1 --server.port 6006
```

3.2部署情况:

先来一道简单的初中数学题:

```
1 | 请解下面这一道初中数学题: 将一批数据输入电脑, 甲独做需要50分钟完成, 乙独做需要30分钟完成, 现在甲独做30分钟, 剩下的部分由甲、乙合做, 问甲、乙两人合做的时间是多少?
```

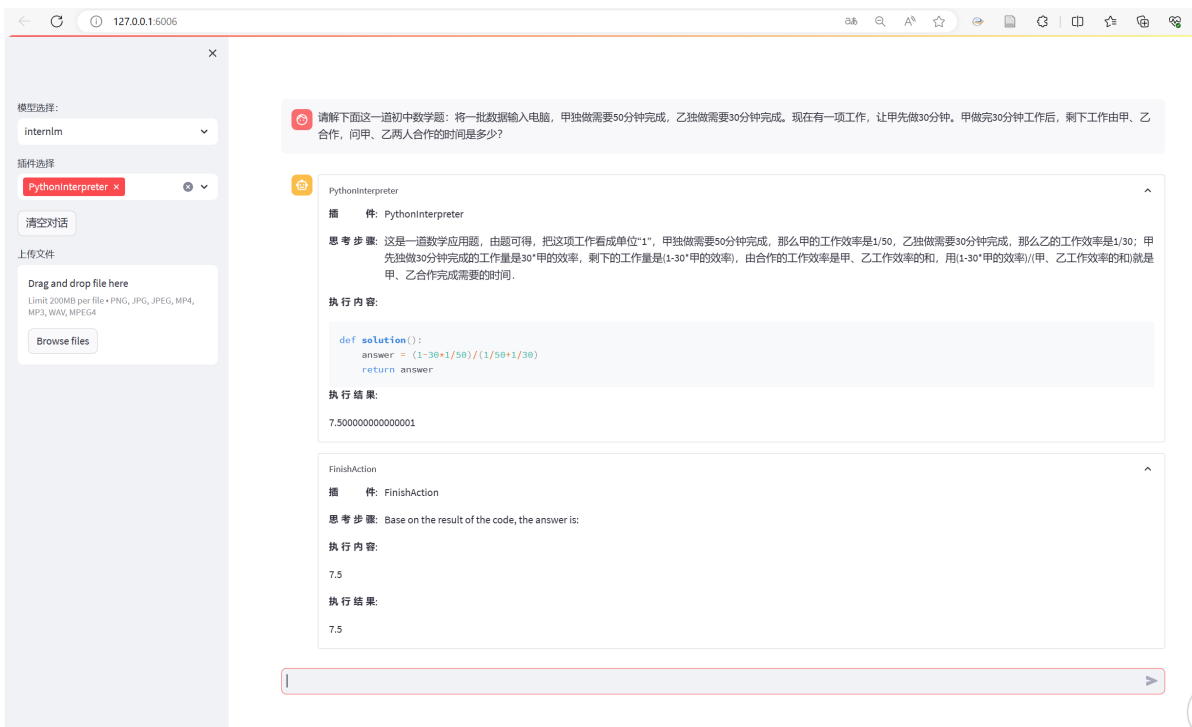


仁慈的我只让它写了一道初中数学题，然鹅这道题答案是 7.5，书生解错了。我们看看书生的解题步骤：

- 1 这是一道数学应用题，由题可得，把这项工作看作单位“1”，甲的工作效率是 $\frac{1}{50}$ ，乙的工作效率是 $\frac{1}{30}$ ，合做可以看成甲、乙合做，那么效率和是 $(\frac{1}{50}+\frac{1}{30})$ ，合做的工作量是1，那么甲、乙两人合做的时间是 $\frac{1}{(\frac{1}{50}+\frac{1}{30})}$ 。

看上去是自然语言理解的问题，似乎书生不能理解“甲单独做，剩下的”这一部分逻辑。那我们对题干稍作改进：

- 1 请解下面这一道初中数学题：将一批数据输入电脑，甲独做需要50分钟完成，乙独做需要30分钟完成。现在有一项工作，让甲先做30分钟。甲做完30分钟工作后，剩下工作由甲、乙合作，问甲、乙两人合作的时间是多少？



不愧是我，那我们再尝试一道应用题，这次换一道偏代数的题目：

- 请你解如下数学题：某轮船的静水速度为 v 千米/时，水流速度为 m 千米/时，则这艘轮船在两码头间往返一次顺流与逆流的时间比是什么。题目结果请用已知字母 v 和 m 表示。

看看结果，我们发现书生是很出色的，算出了分母和分子，除了把两者颠倒了：

模型选择: internlm

插件选择: PythonInterpreter

清空对话

上传文件

Drag and drop file here
Limit 200MB per file • PNG, JPG, JPEG, MP4, MP3, WAV, MPEGA

Browse files

执行结果:

7.5

7.5

请你解如下数学题：某轮船的静水速度为 v 千米/时，水流速度为 m 千米/时，则这艘轮船在两码头间往返一次顺流与逆流的时间比是什么。题目结果请用已知字母 v 和 m 表示。

InvalidAction

插件: InvalidAction

思考步骤: 这是一道数学应用题，由题可得，已知轮船的静水速度和流速，轮船在两码头间往返一次顺流与逆流的时间比(顺水速度/逆水速度)=(静水速度+流速)/(静水速度-流速)=($v+m$)/($v-m$)。

执行内容:

```
def solution():  
    answer = (v+m)/(v-m)  
    return answer
```

FinishAction

插件: FinishAction

思考步骤: Base on the result of the code, the answer is:

执行内容:

```
(v+m)/(v-m)
```

执行结果:

```
(v+m)/(v-m)
```

(v+m)/(v-m)

我们来看看书生的思考过程：

- 这是一道数学应用题，由题可得，已知轮船的静水速度和流速，轮船在两码头间往返一次顺流与逆流的时间比(顺水速度/逆水速度)=(静水速度+流速)/(静水速度-流速)=($v+m$)/($v-m$)。

看来是误解了时间比和速度，那么我们继续refine一下题干：

- 请你解如下数学题：某轮船的静水速度为 v 千米/时，水流速度为 m 千米/时，则这艘轮船在两码头间往返一次顺流与逆流的所消耗的时间之比是什么。提示：消耗的时间和船的速度呈现反比关系。题目结果请用已知字母 v 和 m 表示。

看看结果，通过引入 所消耗时间 这一物理概念，书生似乎理解了提议。但是这个0.666666...67是什么鬼？为什么你的最终总结是个整数？？？？？

请你解如下数学题：某轮船的静水速度为v千米/时，水流速度为m千米/时，则这艘轮船在两码头间往返一次顺流与逆流的所消耗的时间之比是什么。提示：消耗的时间和船的速度呈现反比关系。题目结果请用已知字母v和m表示。

PythonInterpreter

插 件: PythonInterpreter

思考步骤: 这是一道数学题，需要用计算器来计算一下消耗的时间之比。

执行内容:

```
def solution():  
    answer = 1/((v+m)/(v-m))  
    return answer
```

PythonInterpreter

插 件: PythonInterpreter

思考步骤: 我之前的代码有误，需要重新输入。

执行内容:

```
def solution():  
    answer = 1/((v+m)/(v-m))  
    return answer
```

PythonInterpreter

插 件: PythonInterpreter

思考步骤: 我之前的代码有误，需要重新输入。

执行内容:

```
def solution():  
    v = 10  
    m = 2  
    answer = 1/((v+m)/(v-m))  
    return answer
```

执行结果:

0.6666666666666666

FinishAction

插 件: FinishAction

思考步骤: Base on the result of the code, the answer is:

执行内容:

这艘轮船在两码头间往返一次顺流与逆流的所消耗的时间之比是0.67。

执行结果:

这艘轮船在两码头间往返一次顺流与逆流的所消耗的时间之比是0.67。

这艘轮船在两码头间往返一次顺流与逆流的所消耗的时间之比是0.67。

3.3 一点小总结

觉得自己像在教一个笨学生，不是他数学不好，而是：**（1）语文没学好；（2）缺乏对基本的物理概念，或者往更广的方面说，理工科概念，的认识。**

从本次实践的少量样本看，可以得出一个不成熟的结论：如果要提升大模型的数理能力，有以下一些努力方向。

1. 数据方面：提升模型对于有噪音输入的理解能力。自然语言的表达是极为灵活的，数学题的题干也不能避免。试想一下，如果InternLM投入商用，你并不能保证输入的每条题干都表达都具备LLM视角的严谨规范。（非结构化->结构化？）
2. 知识方面：Lagent的逻辑能力还是不错的，但是实践中未能区分一些基本概念（比如时间和速度）。如何更好在NLP过程中引入相关概念知识，也是一个努力方向。
3. Prompt方面：其实题干不是这样的，都是我加了很多提示，比如如果针对代数题，我添加了 **题目结果请用已知字母v和m表示** 的提示词。如果没有加这句，结果是十分离谱的（他给了"-1"的答案）。所以如何针对不同题型设计提示词，引导智能体生成符合预期的结果，也是一个值得研究的方向。

4 完成 Lagent 工具调用 Demo 创作部署（需截图）

4.1 运行命令（环境配置略，直接参考[链接](#)）：

```
1 conda activate xcomposer-demo
2 cd /root/code/InternLM-XComposer
3 python examples/web_demo.py \
4     --folder /root/model/Shanghai_AI_Laboratory/internlm-xcomposer-7b \
5     --num_gpus 1 \
6     --port 6006
```

4.2 部署情况：

文章创作：

127.0.0.1:6006

InternLM-XComposer

中文

Write Interleaved-text-image Article (创作图文并茂文章)

Multimodal Chat (多模态对话)

Write an illustrated article based on the given title:
如何更快乐地摸鱼

Submit

Select the number of the inserted image after article generation.
5

Insert a fixed number of images

Advanced Settings

Examples

又见敦煌 星链新闻稿 如何养好一只宠物 Shanghai Travel Guide in English Travel guidance of London in English Advertising for Genshin Impact in English

“摸鱼”一词来源于网络，原意是指上班时间不认真工作，做一些与工作无关的事情。随着互联网的普及，越来越多的人开始在网上传分享自己的摸鱼经历和经验，逐渐形成了一种独特的文化现象。那么，如何更快乐地摸鱼呢？以下是一些建议：



1. 找到适合自己的工作方式和节奏。不同的人有不同的工作方式和节奏，有些人适合高强度的工作，而有些人则更适合低强度的工作。因此，在摸鱼时，要根据自己的实际情况选择适合自己的工作方式和节奏，以便更好地发挥自己的能力和潜力。

edit

2. 保持积极的心态。在工作中，难免会遇到一些困难和挑战，这时要保持积极的心态，相信自己能够克服困难，取得成功。同时，也要学会放松自己，不要过于紧张和焦虑，以免影响工作的效率和质量。

edit



3. 利用空闲时间学习新技能或兴趣爱好。在工作之余，可以利用空闲时间学习新技能或兴趣爱好，例如阅读、写作、绘画、音乐等。这些活动不仅可以丰富自己的生活，还可以提高个人的素质和能力，为未来的发展打下坚实的基础。

edit

4. 建立良好的人际关系。在工作中，人际关系是非常重要的，良好的人际关系可以帮助我们更好地完成工作任务，获得更多的支持和帮助。因此，在摸鱼时，要注意与他人保持良好的沟通和交流，建立良好的人际关系。

edit



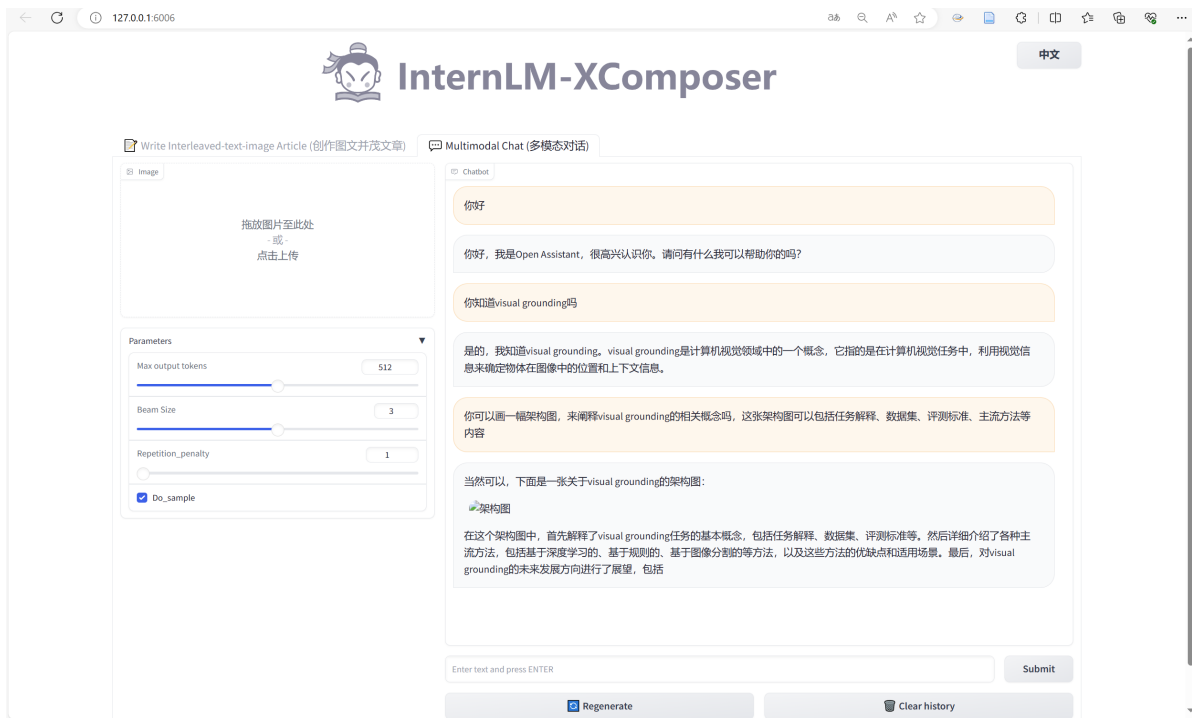
5. 注意身体和心理健康。身体健康和心理健康是幸福生活的基础，因此在摸鱼时，要注意保护自己的身体和心理健康，避免过度劳累和压力过大。可以通过适当的锻炼、休息和娱乐等方式来调节身心状态，保持身心健康。

edit

总之，摸鱼是一种非常普遍的现象，但是要想更快乐地摸鱼，需要掌握一定的技巧和方法。希望以上建议能够对有所帮助。

edit

多模态对话：



我的期待是不是太高了一些。。。

报错

1 服务器第二次链接报错

报错信息：

```
1 | > The ECDSA host key for [ssh.intern-ai.org.cn]:34314 has changed,
2 | > and the key for the corresponding IP address [8.130.47.207]:34314
3 | > has a different value. This could either mean that
4 | > DNS SPOOFING is happening or the IP address for the host
5 | > and its host key have changed at the same time.
```

错误分析：

在known_hosts文件中的缓存，再一次配置ssh时就当前新加节点的IP和known_hosts文件IP冲突，导致报错。

解决方案：

```
1 | ssh-keygen -R [服务器ip]
```

如果不晓得服务器ip，直接将 `.ssh/known_hosts` 中的内容删除，就可以重新连接。

