

第二节 《轻松玩转书生·浦语大模型趣味 Demo》课后作业



https://www.bilibili.com/video/BV1AH4y1H78d/?spm_id_from=333.788&vd_source=628bd23c68972a5700988ed6b195c8e1

部署InternLM2-Chat-1.8B模型：

新建开发环境并配置conda环境。

安装所需的多个库，如huggingface-hub、transformers等，以支持模型运行。

下载并运行InternLM2-Chat-1.8B模型，通过Jupyter terminal进行测试。

部署八戒-Chat-1.8B模型：

切换到已配置的demo环境，从仓库下载模型。

运行模型并通过端口映射使其可以在本地浏览器上访问，熟悉整个操作过程。

运行InternLM2-Chat-7B模型的智能体Lagent：

需要特定的系统权限和等待时间，设置软连接指向模型存放路径。

启动智能体模型，并在本地端口尝试使用，包括解决复杂问题如鸡兔同笼问题。

部署浦语·灵笔2模型：

安装所需的环境包，下载并检出特定版本的源代码。

启动模型，准备模型以供最终使用。

已关注 Tutorial/helloworld/hello.py x +

常用网址 联邦学习书签 差分隐私书签 密码学书签 算法练习书签

打开 lagent 路径:

```
cd /root/demo/lagent
```

在 terminal 中输入指令, 构造软链接快捷访问方式:

```
ln -s /root/share/new_models/Shanghai_AI_Laborator
```

打开 lagent 路径下 examples/
internlm2_agent_web_demo_hf.py 文件, 并修改对应位置 (71
行左右) 代码:

```
# 其他代码...  
value='/root/models/internlm2-chat-7b'  
# 其他代码...
```

输入运行命令 - 点开 6006 链接后, 大约需要 5 分钟完成模型
加载:

创建开发机 WebIDE

https://studio.intern-ai.org.cn/console/trinity/20240330-7a3b4c

常用网址 联邦学习书签 差分隐私书签 密码学书签 算法练习书签 论文阅读 LLM+Fuzzing Le Chat Mistral Github Portfolio me KOOK

CPU: 4.76% GPU: 30% Mem: 4100 MB 内存: 5.52 / 72 GB 0.74% 显存: 9 / 24596 MB 0%

File Edit View Run Kernel Tabs Settings Help

Filter files by name

File Explorer

Name	Last Modified
internlm2_agent_cli_demo.py	5 minutes ago
internlm2_agent_web_demo_hf.py	5 minutes ago
internlm2_agent_web_demo.py	5 minutes ago
model_cli_demo.py	5 minutes ago

```
st.session_state['model_path'] = None  
  
def setup_sidebar(self):  
    """Setup the sidebar for model and plugin selection."""  
    # model_name = st.sidebar.selectbox("模型选择", options=['internlm'])  
    model_name = st.sidebar.text_input("模型名称:", value='internlm2-chat-7b')  
    meta_prompt = st.sidebar.text_area("系统提示词", value=META_CN)  
    da_prompt = st.sidebar.text_area("数据提示词", value=INTERPRETER_CN)  
    plugin_prompt = st.sidebar.text_area("插件提示词", value=PLUGIN_CN)  
    model_path = st.sidebar.text_input("模型路径:", value='internlm/internlm2-chat-20b')  
    if model_name != st.session_state['model_selected'] or st.session_state['  
        'model_path'] != model_path:  
        st.session_state['model_path'] = model_path  
        model = self.init_model(model_name, model_path)  
        self.session_state.clear_state()  
        st.session_state['model_selected'] = model_name  
        if 'chatbot' in st.session_state:  
            del st.session_state['chatbot']  
    else:  
        model = st.session_state['model_map'][model_name]  
  
    plugin_name = st.sidebar.multiselect("插件选择",  
        options=list(st.session_state['plugin_map'].keys()),  
        default=[])  
    da_flag = st.sidebar.checkbox("数据提示",  
        value=False,  
    )  
    plugin_action = [  
        st.session_state['plugin_map'][name] for name in plugin_name  
    ]  
  
    if 'chatbot' in st.session_state:  
        if len(plugin_action) > 0:  
            st.session_state['chatbot']._action_executor = ActionExecutor(  
                actions=plugin_action  
            )  
        else:  
            st.session_state['chatbot']._action_executor = None  
        if da_flag:  
            st.session_state['chatbot']._interpreter_executor = ActionExecutor(  
                actions=[PythonInterpreter()]  
            )  
        else:  
            st.session_state['chatbot']._interpreter_executor = None  
        st.session_state['chatbot']._protocol._meta_template = meta_prompt  
        st.session_state['chatbot']._protocol.plugin_prompt = plugin_prompt  
        st.session_state['chatbot']._protocol.interpreter_prompt = da_prompt  
        if st.sidebar.button("清空对话", key='clear'):  
            self.session_state.clear_state()  
        uploaded_file = st.sidebar.file_uploader("上传文件")  
  
    return model_name, model, plugin_action, uploaded_file, model_path  
  
def init_model(self, model_name, path):  
    """Initialize the model based on the input model name."""  
    st.session_state['model_map'][model_name] = HFTransformer(  
        path=path,  
    )
```

Simple 1 Python

1080P 高码率 倍速

Ln 1, Col 1 Spaces: 4 internlm2_agent_web_demo_hf.py

10:38 / 21:14

搜索

1009 2024/3/30

