ML-Platform使用文档

部署

构建

前端需要使用yarn进行构建,需要npm 16,更高的版本可能有支持问题。输入如下指令安装yarn

```
npm install -g yarn
```

来到Frontend/ml-platform下,输入命令

```
yarn install
yarn build
```

安装依赖项,并进行编译。编译完的前端在dist文件夹下。

部署

打开Frontend/ml-platform/deploy/nginx.conf

依使用需求,将其中的47.94.221.102替换为部署端的公网ip或局域网ip。

随后返回项目根目录。安装好docker和docker-compose后,输入以下命令

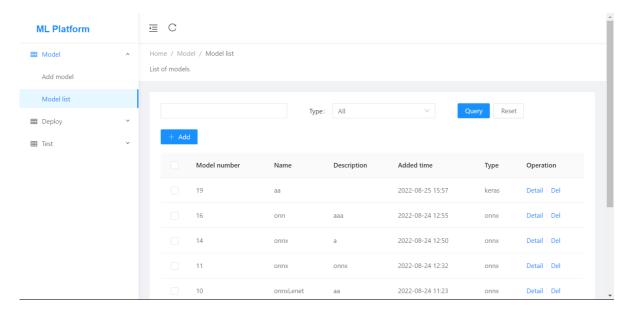
```
docker-compose build
docker-compose up
```

此时已经成功部署,前端在本机8000端口,后端在本机8001端口。

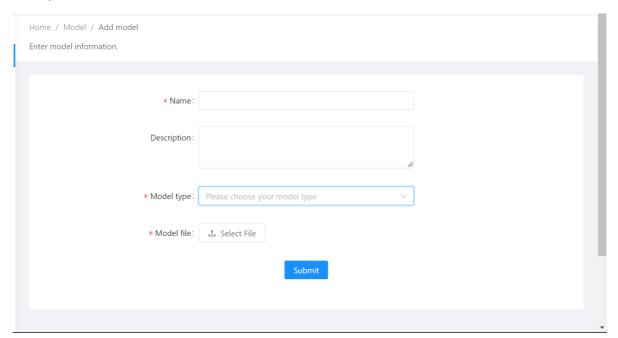
使用

导入模型

进入网站后,点击AddModel或Add按钮,进入导入模型表格

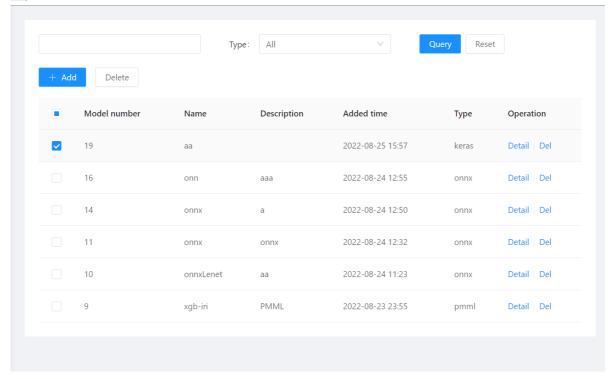


在表格中填写模型名称、模型描述(可选),选择模型类型,上传模型文件(keras模型将文件夹需要打包成zip)。点击Submit进行提交,如果模型不合法,会给出信息。上传完毕后来到模型列表页。



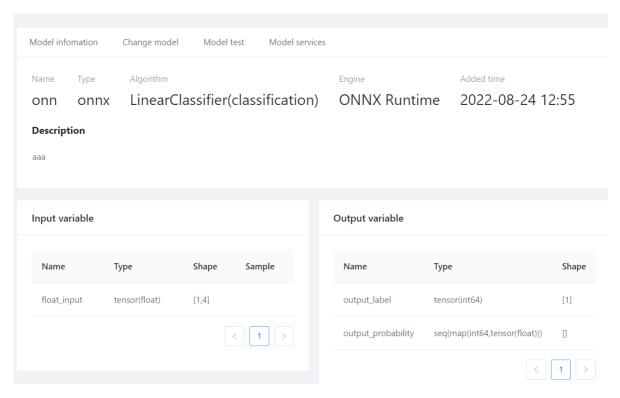
模型列表

可以在搜索框中输入文字,查询会返回名称或描述中包含这些文字的结果。点击Reset清空查询参数。也可以通过模型类型筛选结果。左侧选项可以对模型进行批量删除。点击最右侧连接访问详情或删除模型。

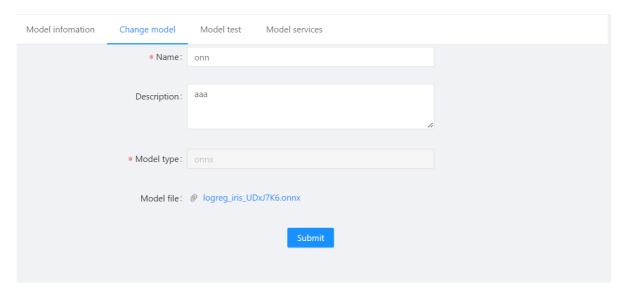


模型详情

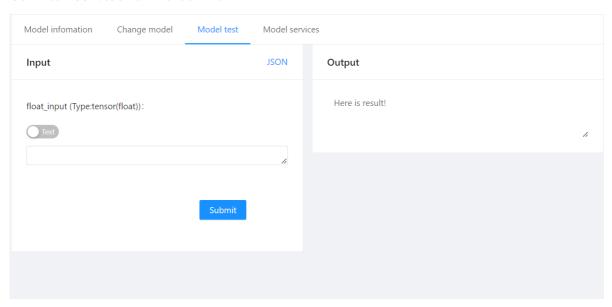
详情页会显示模型信息(名称、类型、算法、引擎、输入、输出、添加时间),如有,则会显示模型描述

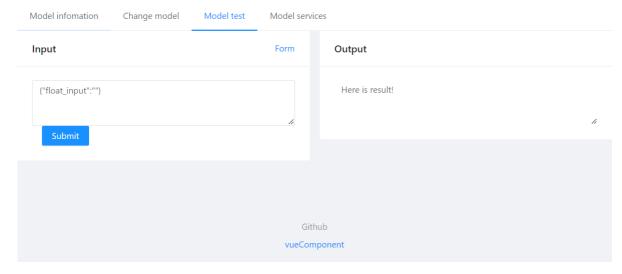


点击Change model,可以修改模型名称、描述。点击链接下载模型文件

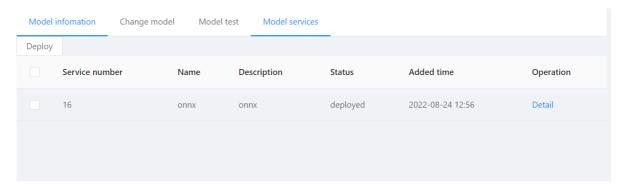


点击Model test,可以对模型进行测试。输入接受表单与JSON两种格式,其中表单各项可以是数据也可以是文件。测试结果会展示在右边的框里。

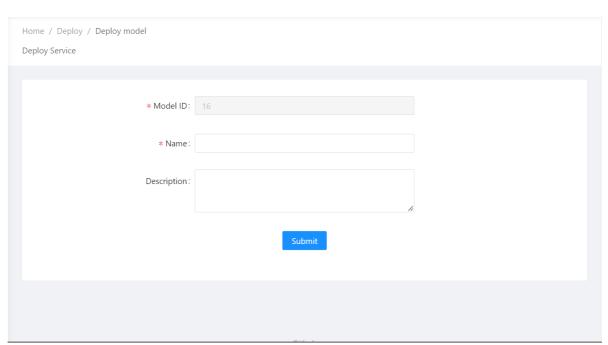




点击Model service查看目前模型部署的服务。点击Detail查看服务详情,点击Deploy对模型进行部署。



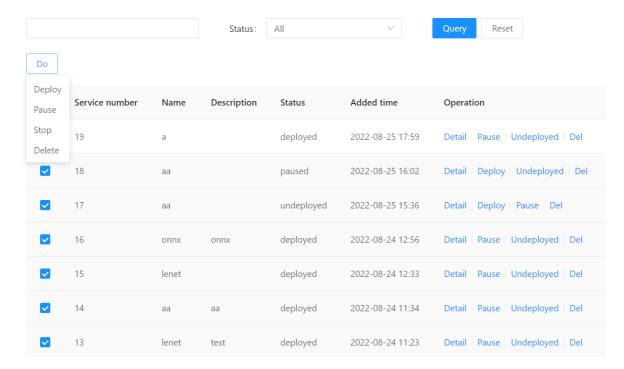
部署模型



输入服务名称、服务描述 (可选),点击Submit部署模型。

服务列表

可以在搜索框中输入文字,查询会返回名称或描述中包含这些文字的结果。也可以通过状态筛选结果。 点击Reset清空查询参数。左侧选项可以对服务进行批量处理。点击最右侧连接访问详情或删除或修改状态。

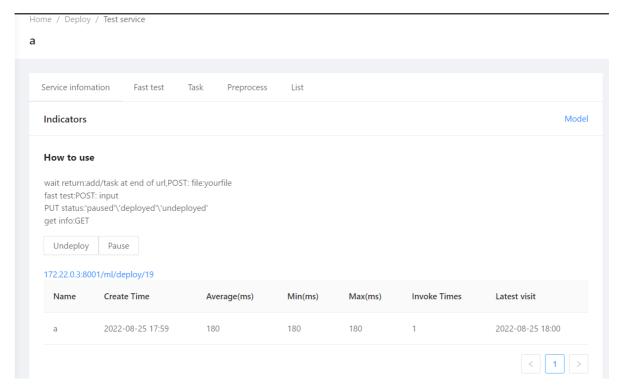


服务详情

详情页会显示服务信息(名称、平均任务使用时间、最大任务时间、最小任务时间、调用次数、添加时间),如有,则会显示服务描述。

点击Model访问服务对应的模型页。

为服务提供了Restful Api接口(注意显示的ip是局域网ip,可以根据需求转为公网ip)



以图中为例

对<u>http://172.22.0.3:8001/ml/deploy/19</u>使用GET请求获取服务信息。

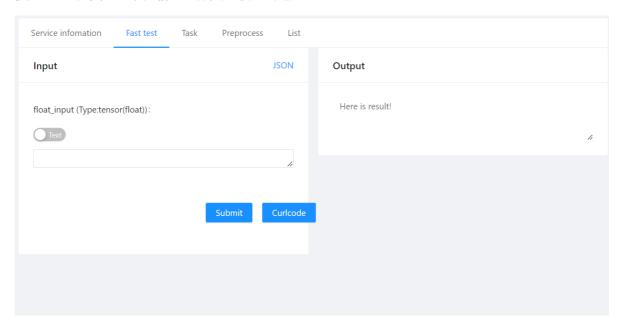
使用POST请求,数据各项为模型输入各项名称及内容,是快速返回接口。

使用PUT请求,数据中含status项,内容为paused或deployed或undeployed,修改服务状态为暂停、部署、停止。

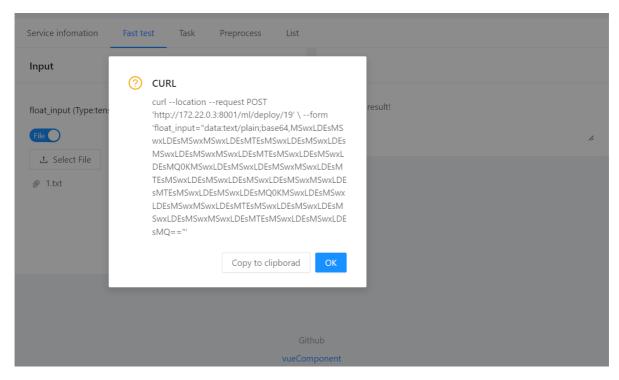
使用DELETE请求,是删除服务接口。

对<u>http://172.22.0.3:8001/ml/deploy/19/task</u>

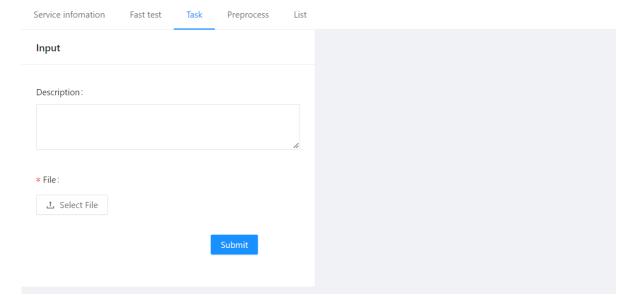
使用POST请求, file项为输入文件, 是等待返回接口。



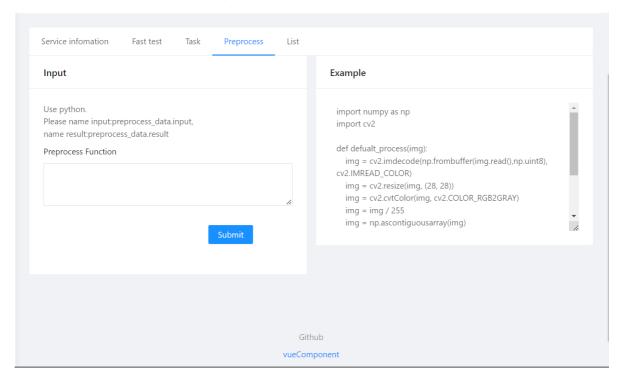
点击Fast test,是快速返回测试。接受表单与JSON,输入完毕后点击Curlcode可以生成curl代码,文件会转为base64。可以复制到剪贴板。根据需求可以将此处ip由局域网ip改为公网ip。



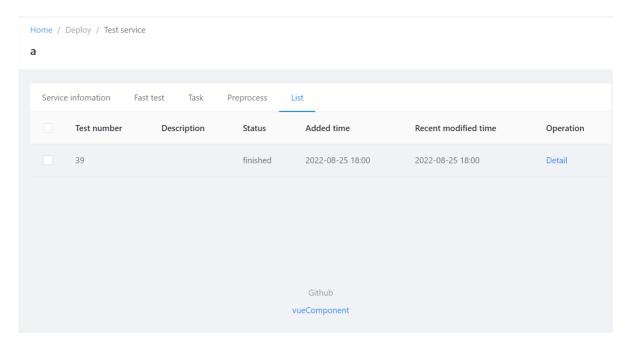
点击Task,是等待返回接口。上传文件,输入描述(可选),会返回任务id,跳转到任务界面。



点击Preprocess, 进入自定义预处理api, 右边是示例。接受python脚本,输入为preprocess_data.input, 输出为preprocess_data.result。可以在快速返回接口进行调试

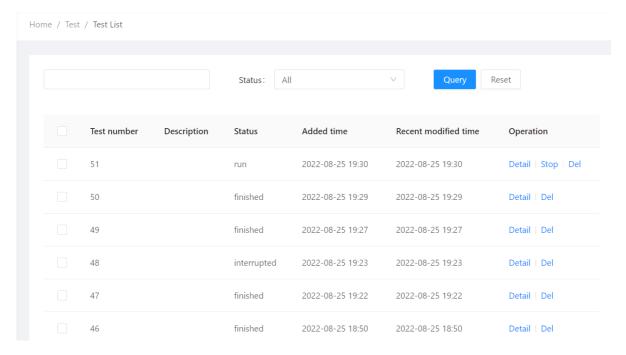


点击List查看目前服务进行的任务。点击Detail查看任务详情。

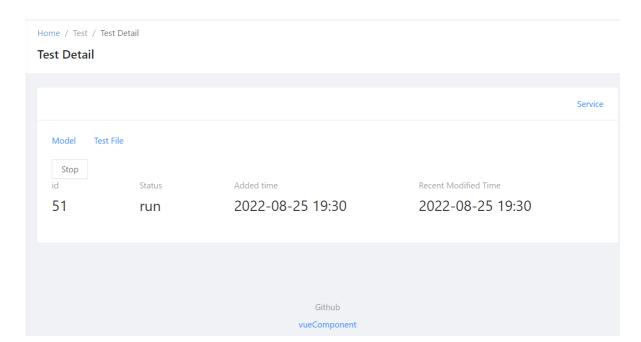


任务列表

可以在搜索框中输入文字,查询会返回名称或描述中包含这些文字的结果。也可以通过状态筛选结果。 点击Reset清空查询参数。左侧选项可以对任务进行批量处理。点击最右侧连接访问详情或删除或修改状态。



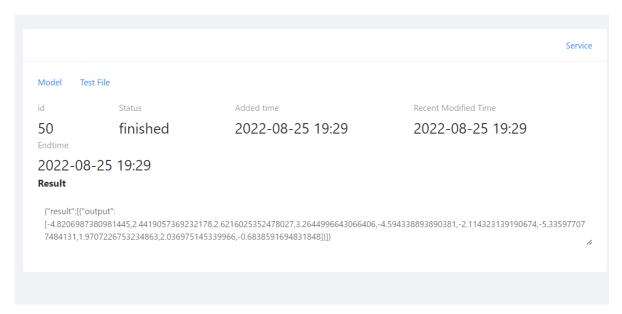
任务详情



未完成的任务可以点击Stop中断。点击Service返回服务页,点击Model来到服务使用的模型的模型页,点击Test File下载任务输入的文件。

完成的任务会显示完成时间与结果。

Test Detail



因报错而被打断的任务会显示报错。

