

Gradio-HuggingFace 的实践探索

参考资料:

- gradio-GitHub: [Link-1](#)
Create UIs for your machine learning model in Python in 3 minutes
- Colab 实践: [Link-2](#)

Introduction

训练完机器学习模型后，接下来要来做的是通过演示向全世界展示它。目前最简单的方法是使用Gradio，托管到HuggingFace Spaces上。在Spaces上部署Gradio架构后，模型部署只需要不到10分钟！让我们看看如何轻松地向全世界部署第一个模型来使用这个平台。

Using Gradio

```
# 1. load the model
learn = load_learner('export.pkl')

# 2. Define a prediction function model
labels = learn.dls.vocab
def predict(img):
    img = PILImage.create(img)
    pred, pred_idx, probs = learn.predict(img)
    return {labels[i]: float(probs[i]) for i in range(len(labels))}

# 3. 最后，让我们导入Gradio，利用它的功能制作一个界面并启动它。注意，如果你是在笔记本上做的，
Gradio的演示也会在笔记本上显示出来，供你互动尝试（这里我只显示截图）。
# hide_output
import gradio as gr
gr.Interface(fn=predict, inputs=gr.inputs.Image(shape=(512, 512)),
            outputs=gr.outputs.Label(num_top_classes=3)).launch(share=True)
```

Gradio的内网部署实践

目前遇到的主要问题:

- 无法对默认设置的127.0.0.1的地址进行访问;

测试案例:

```
import gradio as gr
from gradio import Interface

def greet(name):
    return "Hello" + name + "!"

demo = gr.Interface(fn=greet, inputs='text', outputs='text')
demo.launch()
```

```
python test_gradio.py
```

重置Server服务器后的设置:

```
import gradio as gr
from gradio import Interface

def greet(name):
    return "Hello" + name + "!"

demo = gr.Interface(fn=greet, inputs='text', outputs='text')
demo.launch(server_name='0.0.0.0')
```

```
python test_gradio.py
```

遇到的问题分析

1. ImportError: cannot import name 'Interface' from 'gradio'
(/home/jovyan/work/gradio.py)

问题分析: 文件名的命名问题, 需要对文件进行重命名: `gradio.py --> test_gradio.py`

NOTE: 该问题已经出现多次, 需要格外注意。

2. 开放的端口无法正常访问:

内网ip 地址: 10.17.68.105 : 7862, 容器内部的放开端口为: 127.0.0.1:7860-7865 --> 7860-7865/tcp

报错: This site can't be reached.

原因分析:

- server的设置问题, `gradio`默认使用127.0.0.1作为服务地址, 从而导致外部接口无法进行访问, 需要将server设置为'0.0.0.0'。
- 参考:
`gradio-app/gradio: Issues`