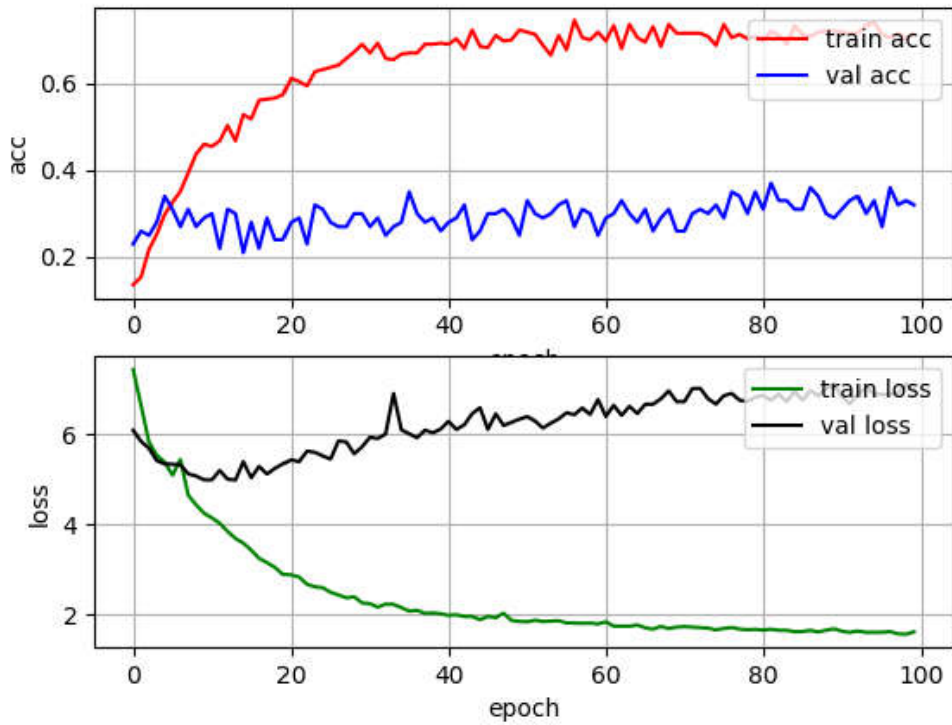


# 周报

这周把实验代码的控制部分改完了，目前的实验结果如图：



训练数据有收敛，准确率能达到 71%，但是验证集的效果还是太差了，原因可能是：

- 1.由于训练数据太少，不能较好的覆盖
- 2.有些标签太少，数据中只出现了几次
- 3.不知道是不是代码还有问题

attention

目前实验中，记忆模型的空间定为 5 轮，查看其 attention 发现，有 80%的相关数据能被模型注意到，如：

I checked the availability for this date and there were no trips available. Would you like to select some alternate dates?

I checked the availability for that date and there were no trips available. Would you like to select some alternate dates?

这两句的注意力得分很高

但是，对于省略的表达，注意力得分不高

Would you be interested in Calgary?

that would be nice

这个用目前的注意力机制很难发现，that would be nice 是指对 Calgary 感兴趣。这种利用基于相关性的 attention 机制很难去处理。

## 控制门

目前的控制门，感觉还是存在一些问题，如果历史记忆较少，那么控制门的  $\alpha$  较大，表明模型会较大可能的记住历史信息，帮助分类。而如果历史数据较多，那么控制门的  $\alpha$  几乎为 0,表明模型会尽可能的忘记历史信息。这个和期望中的还是不太一样，因为历史记忆相关和不相关，对于控制门的影响偏小了。

遇到的问题：

之前对于如何获取中间层的输出？[查看官网](#)

## 如何获取中间层的输出？

一种简单的方法是创建一个新的 `Model`，使得它的输出是你想要的那个输出

```
from keras.models import Model

model = ... # create the original model

layer_name = 'my_layer'
intermediate_layer_model = Model(input=model.input,
                                  output=model.get_layer(layer_name).output)
intermediate_output = intermediate_layer_model.predict(data)
```

此外，我们也可以建立一个Keras的函数来达到这一目的：

```
from keras import backend as K

# with a Sequential model
get_3rd_layer_output = K.function([model.layers[0].input],
                                   [model.layers[3].output])
layer_output = get_3rd_layer_output([X])[0]
```

当然，我们也可以直接编写Theano和TensorFlow的函数来完成这件事

但是这个仅仅适用于一个层只与一个输入相连，之前实现的模型都是单个输入，没有发现这个问题。

但如果一个层有多个输入，就会有问题：

```
D:\360data\Anaconda3\python.exe D:/file/intent_model/SLU/code/test.py
Using TensorFlow backend.
Traceback (most recent call last):
  File "D:/file/intent_model/SLU/code/test.py", line 304, in <module>
    print(lstm.output)
  File "D:\360data\Anaconda3\lib\site-packages\keras\engine\topology.py", line 955, in output
    ' has multiple inbound nodes, '
AttributeError: Layer simple_rnn_1 has multiple inbound nodes, hence the notion of "layer output" is not applicable.

Process finished with exit code 1
```

找了很多方法都不适应，查看源码也是云里雾里，后来通过询问众多大神解决了。

目前的问题是：

之前的代码写的过于粗糙，Attention 的数值计算没有 softmax 处理，导致有些数值太大了。这个可能也导致了历史记忆太多时，控制门数

值很小的原因，导致和不想关的历史记忆不能区分了，这个我正在修改代码，看看是否效果会有所改进。