

1. Model

Source code 主要參考了 tensorflow 官網上的 seq2seq 的 model。

我們使用了以 GRU 為主的有 attention 機制的 encoder-decoder model。在個別的 task 的共同設定兩個語言的 vocabulary 皆為 40000，使用 batch size 64 以及三層的 seq2seq model。在 translation 中設定 neural size 為 128，而在 generation 中則設定成 300。

2. 學到什麼以及怎麼 improve

這一份作業主要學到要怎麼在 tensorflow 裡面處理 sequence 的 batch 的處理方法。另外也有要怎麼在 sequence 的 model 裡面使用 attention 機制。而在比賽中可以做的事情就是把 neural size 變大以及讓 training 時間更久一些就會 improve 了。在 generation 的 task 中，我們直接將 json 的格式變換成和 translation 一樣，並且沒有做任何的語言處理，就可以過 strong baseline 了，讓人驚艷 rnn 的強大。

3. 解釋 code 的主要函式

```
def read_data
```

這主要是把資料 load 進來並且轉換成相對應 vocabulary 的 id 後塞入相對應的 bucket。

```
def create_model
```

這主要是呼叫 tensorflow 中 seq2seq_model 的 api 把 model 初始化的地方。如果在 model 資料夾已經有 trained model 的存在那就會把那個 model load 進來。

```
def train():
```

這主要是 training 的地方，在這裡會把 data 讀進來並且 create model 就開始主要的訓練。

```
def decode():
```

這裡是 training 好後作為吃資料然後預測出結果並且將其儲存的地方。