

MLDS HW2 - Video Captioning

學號：B04901066 系級：電機三 姓名：洪國曉

1. Model description (2%)

使用keras實作，input=(1450, 80+15(輸出句子長度), 4096) ,

兩層LSTM(512) , Dropout(0.25)

output=(1450, 80+15 , 3221(字典大小)) (activation='softmax')

因為觀察發現輸出句子通常很短，因此未實作將上個timestep output回授至下個LSTM輸入，影響為training時間變長。

2. Attention mechanism(2%)

---How do you implement attention mechanism? (1%)

使用github code：

https://github.com/datalogue/keras-attention/blob/master/models/custom_recurrents.py

複製hidden state到序列，然後以 權重矩陣 乘上 重複的隱藏狀態，計算attention probabilities，再求出context vector。接著更新狀態：計算"r"、"z"gate，r-gate決定是否記憶狀態，z-gate決定step的幅度，求得proposal hidden state，最後依此更新狀態即可。

---Compare and analyze the results of models with and without attention mechanism. (1%)

正在跑...

3. How to improve your performance (1%)

---Write down the method that makes you outstanding

---Describe the model or technique (0.5%)

---Why do you use it (0.5%)

一開始字典直接使用單字(即將caption以空格分開)，訓練出來的模型輸出一直是 'A man is a is a a a.'，'A woman is.'...之類的句子，因此我決定在字典動手腳，字典本來存單字，改成存詞組甚至片語，採用以下演算法對caption做前處理：

取代逗號、句號為空白

取代" a "為" a_"，取代" an "為" an_"，取代" the "為" the_"，

取代" one "為" one_"，取代" two "為" two_"，取代" three "為" three_"，

取代" some "為" some_"，取代" there is "為" there_is "，取代" there are "為" there_are "

，

取代" is "為" is_"，取代" are "為" are_"

最後以空格切出詞組放入字典即可。

如此訓練出來的模型BLEU@1至少是2.0起跳。

因為未實作output回授，所以也未實作Schedule Sampling和Beamsearch

4. Experimental results and settings (1%)