MLDS Homewor2 report

1. **group information**

隊伍名稱: 怕～了吧～～

成員: B01901085 楊承翰

B01901121 李律慈

分工:

楊承翰: 資料處理、改善RNN、output處理

李律慈: RNN基本架構、Rmsprop

1. **What we have done**

完成RNN基本架構與Rmsprp，由於沒有可用的GPU，因此這次未實做batch。將training data切成小檔案，分別讀入，但訓練效果不佳。而後初始化使用0.01\* Identity matrix，並限制gradients大小，最終將gradients大小限制在正負0.01，但若將每個epoch跑過所有的train data依然效果不佳，看不到顯著的上升。將每個epoch跑過的data數降低，只取用部分data，有些則不使用，得到較顯著的效果。

而後將output做trimming，並且開一個大小為5的window，window內出現次數在三次以上的才輸出，其餘的不管，將上傳的cost降低了8.9。

1. **Experiment results**
2. 我們一開始先比較cost function使用平方差和softmax的差別，發現使用softmax明顯收斂得比較快。

Number of data = 20

Number of data = 100

1. 接著我們增加training data個數，發現只使用gradient descend 的話，經過2000次iteration仍然卡在local minimum(正確率約0.19)，而加上rmsprop後，發現cost震動的比較劇烈，但是在約第300次iteration之後就跳脫local minimum了。
2. 最後，由於怕訓練時間不好掌控，我們使用了漸進增加training data的方式，每次正確率達到一定值之後，先用目前的model產出依次結果，接著將data數量加倍繼續訓練，得到不錯的結果。