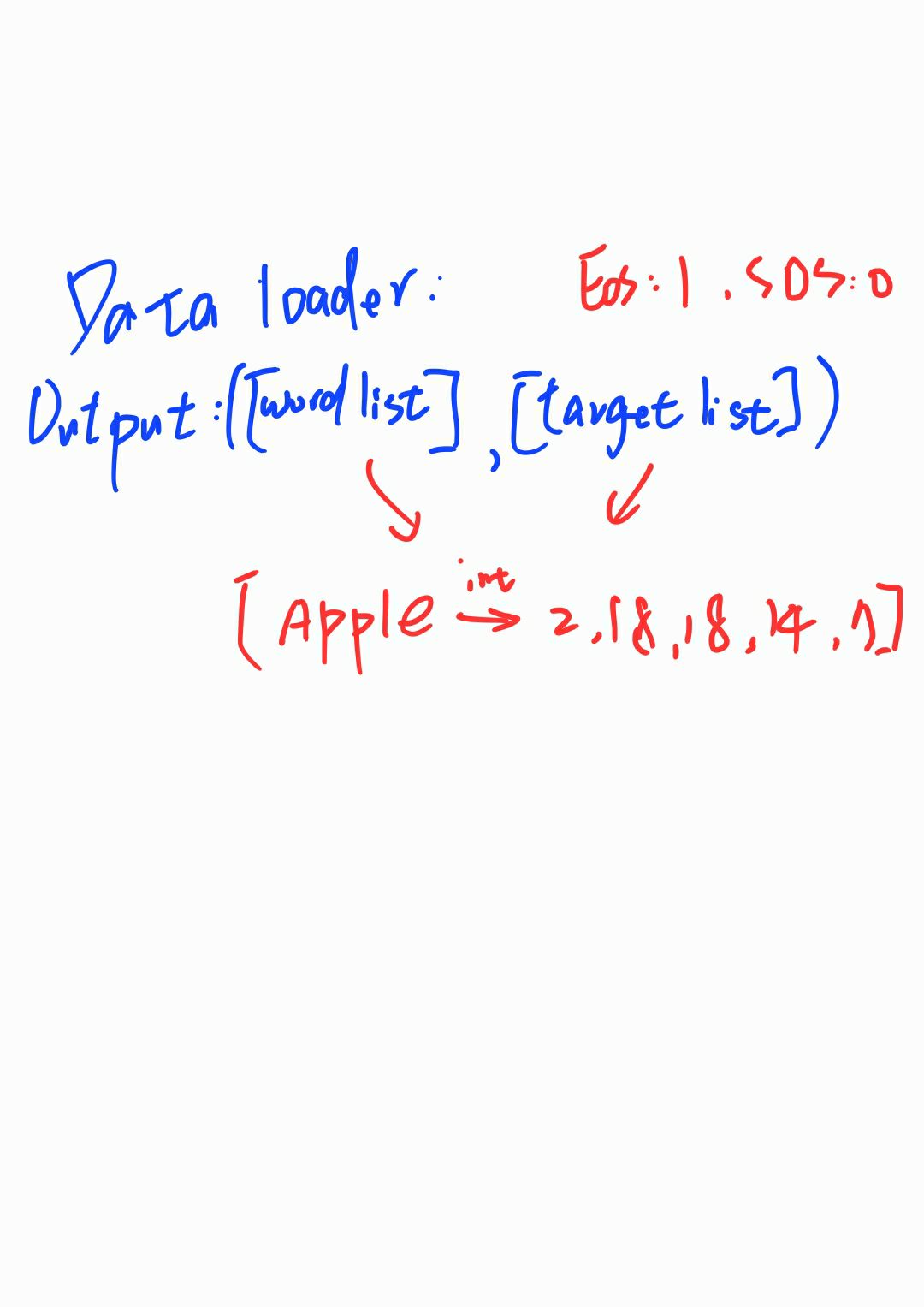
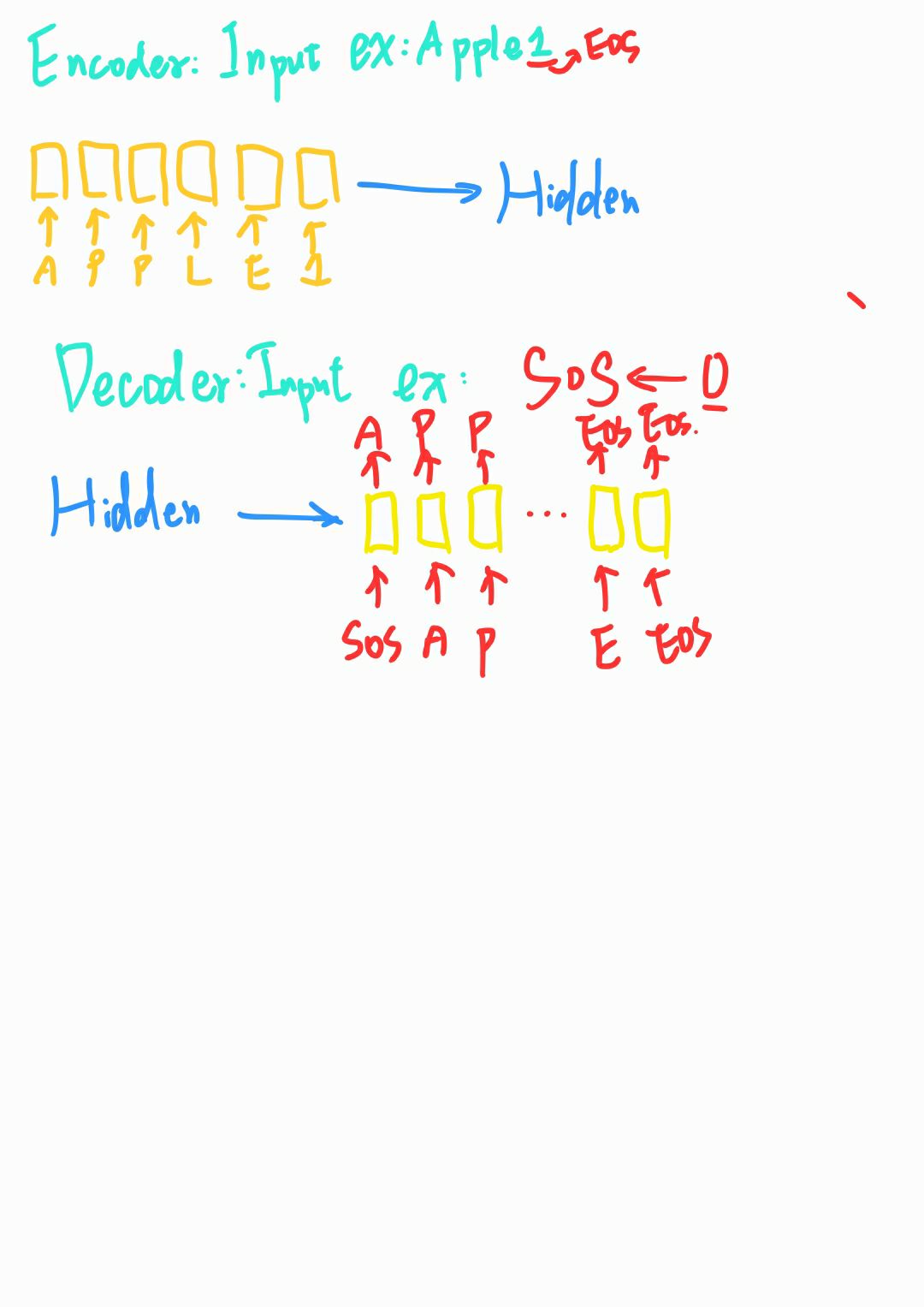
Evalutioan : bleu\_score, word\_acc(整個單字全對才算分)

Simple LSTM :

Data structure:



LSTM model:



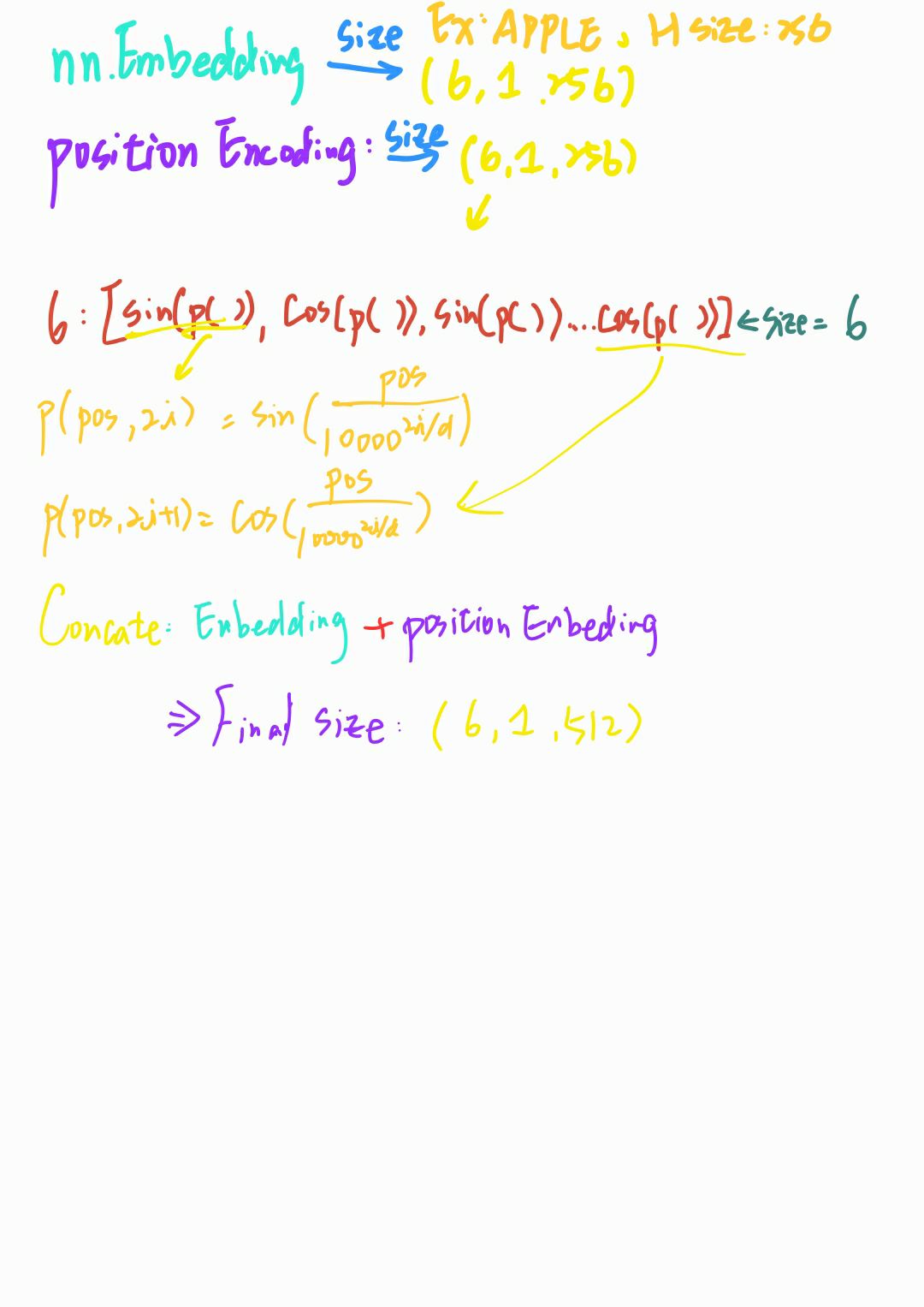
本次功課我測試了最陽春的LSTM(hidden layer : 256 ,num layer: 1)來訓練。

發現只要訓練6,70 epoches，在test, new\_test上分數的表現分別可以達到0.86,0.63

一張含有 文字 的圖片

自動產生的描述

以上面的simple LSTM為基準，我另外加上了position encoding。

實際上為implement為下圖這個樣子。

實際上程式碼:

一張含有 文字 的圖片

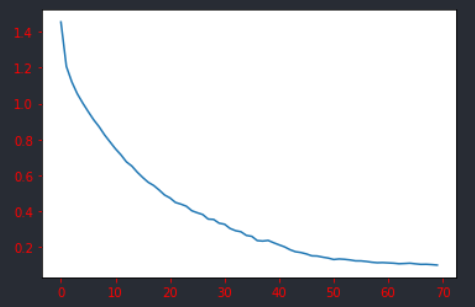
自動產生的描述

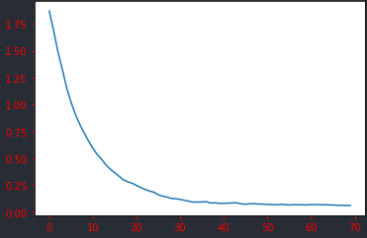
最後發現 : 加了position encoding後，

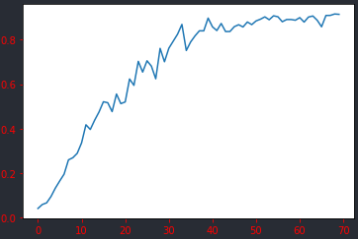
1. 模型收斂速度變快。(simple : 60epochs, position encoding : 40epochs)
2. 在test上score的表現差不多。(simple : 90, position encoding : 93)
3. 在new test上score的表現也差不多(simple : 0.75, position encoding : 0.73)

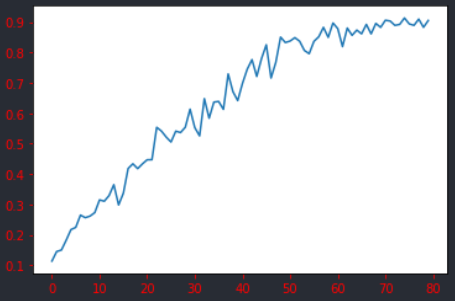
Plot :

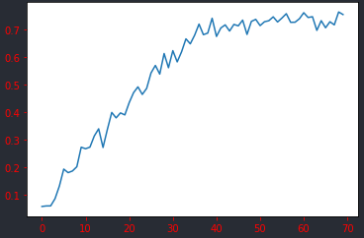
simple position encoding

Loss loss



test score test score



New\_test score New\_test score

