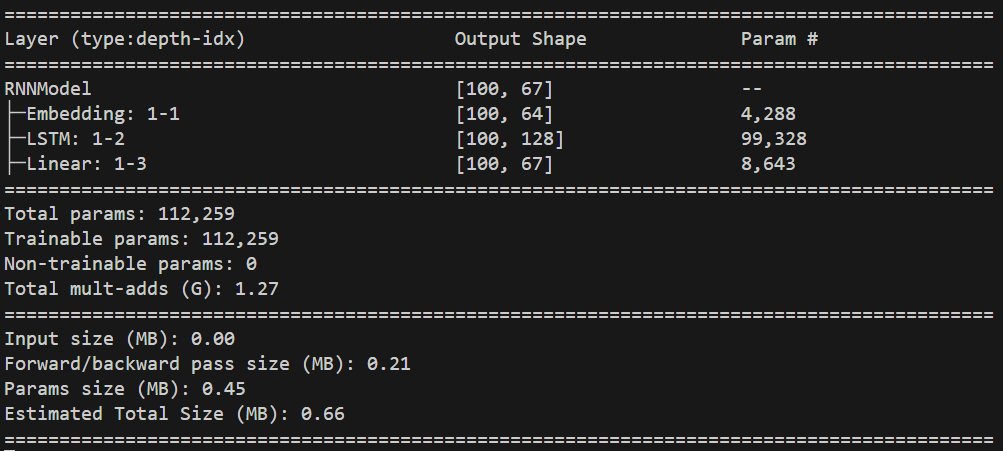
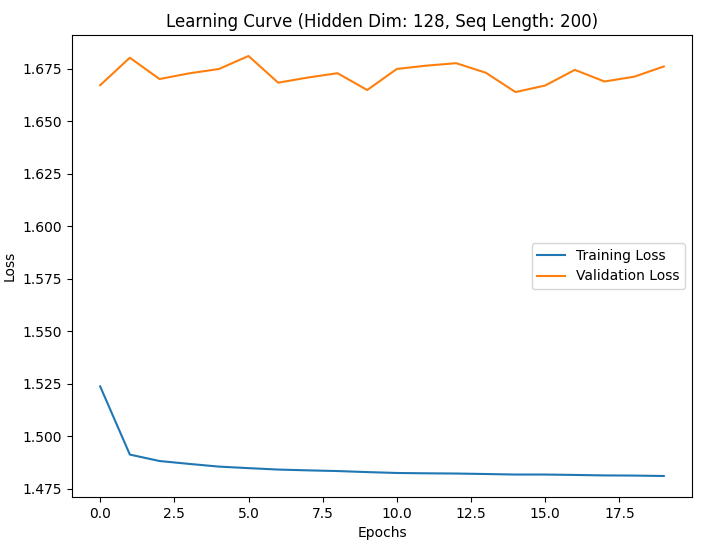
Deep Learning -HW3 report 312512011 李效賢

1-1: Simple (Standard) RNN architecture(函數名稱沒改掉不小心弄成LSTM)



1-2: Simple (Standard) RNN learning curve

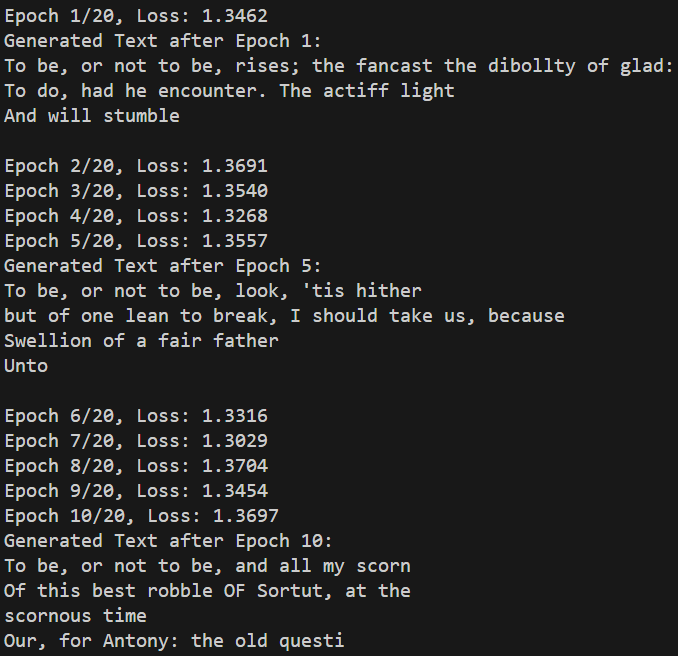


2. 5 breakpoints during training process

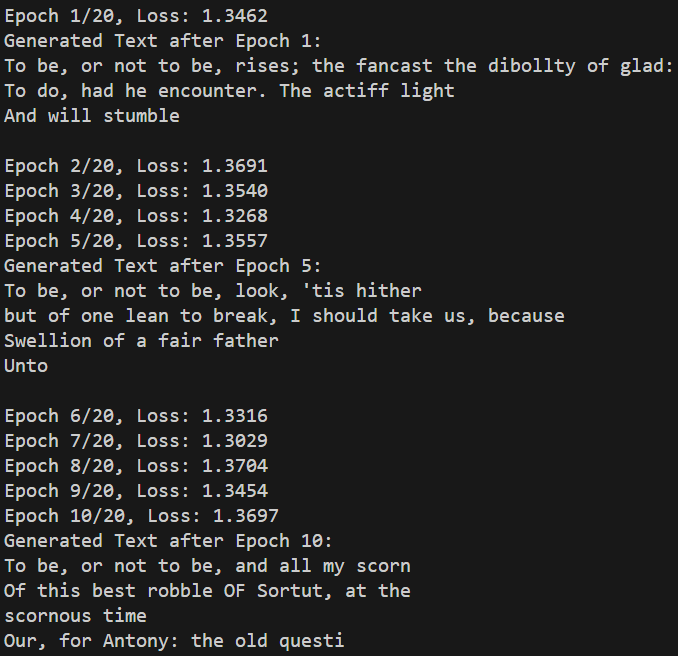
Epoch 1  
一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

Epoch 5



Epoch 10



Epoch 15

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

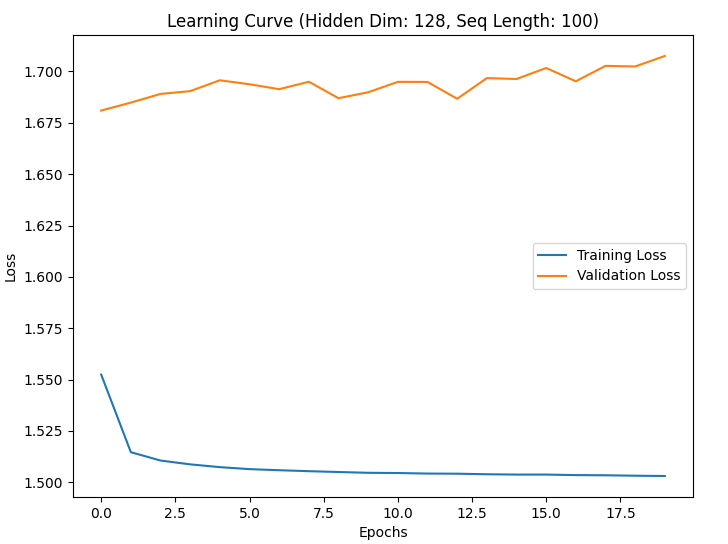
3. Compare different size of hidden states and sequence length

Hidden states = 64, sequence length =100, total parameters = 16963

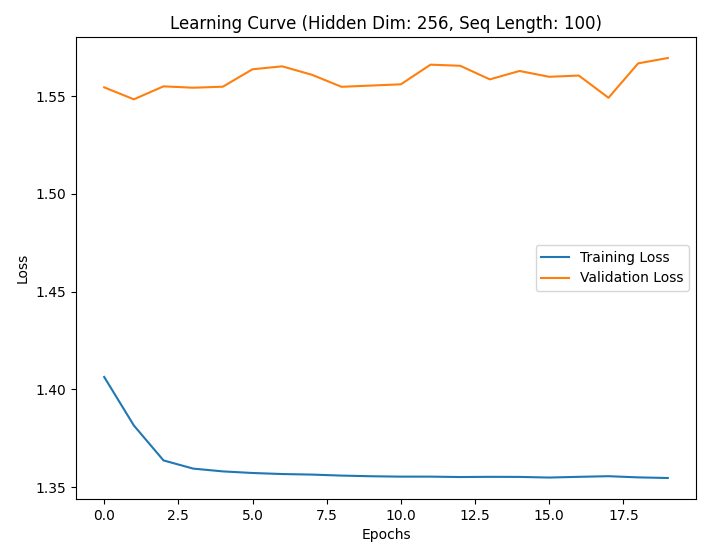
一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片

自動產生的描述

Hidden states = 128, sequence length =100, total parameters = 37763



Hidden states = 256 , sequence length =100, total parameters = 37763

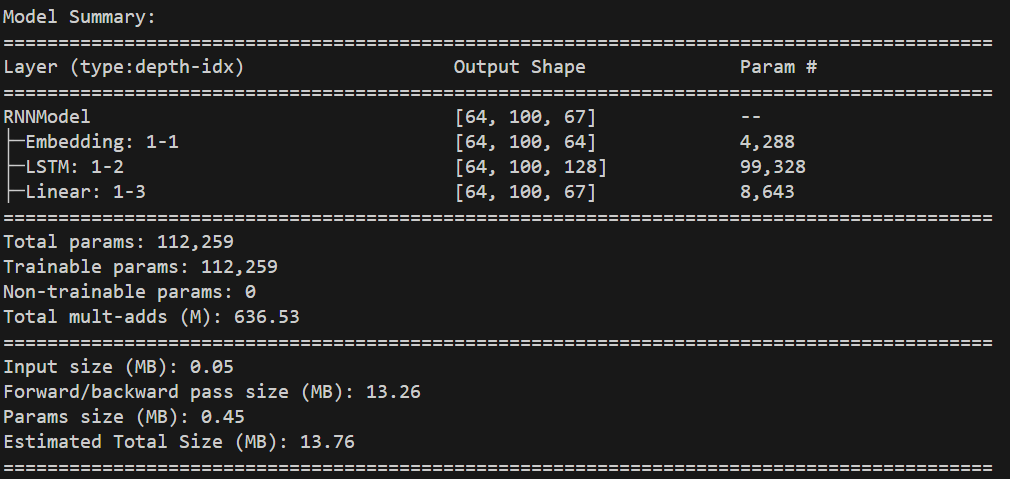


隨著hidden dimension變多，Loss function的值會較小(不論在train還是在

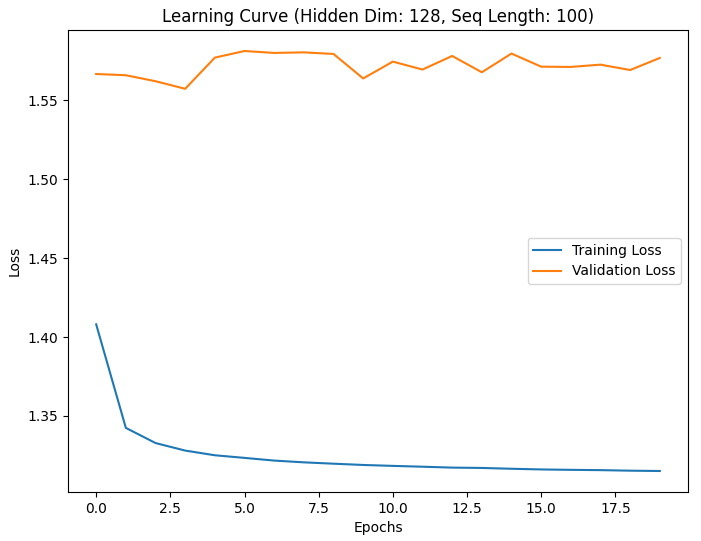
valid的時候都一樣)，此時要小心的是overfitting，但幾乎在訓練5次後loss

function就差不多下降到固值不再有明顯變化。

4-1-1 LSTM network architecture

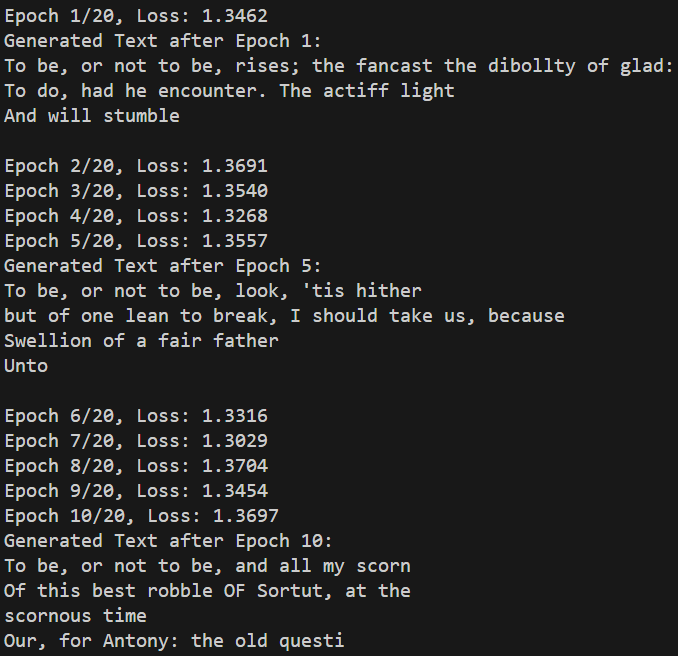


4-1-2 LSTM learning curve

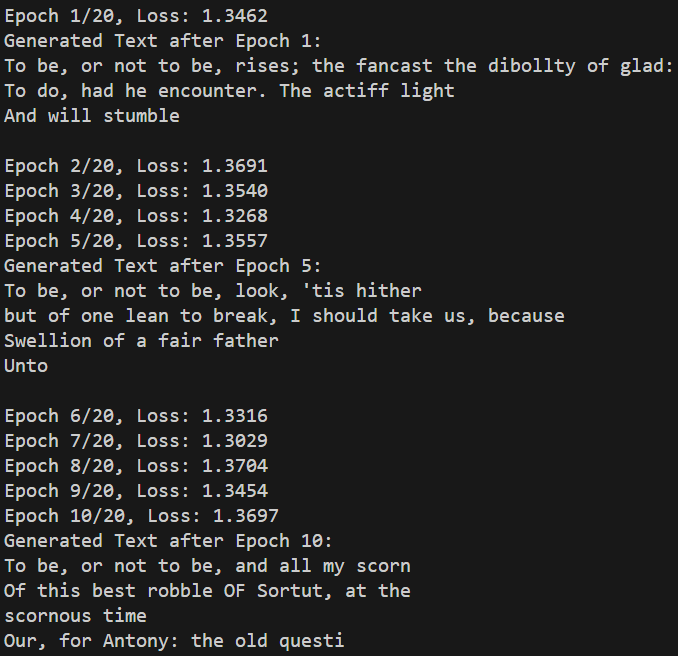


4-2 5 breakpoints during training process

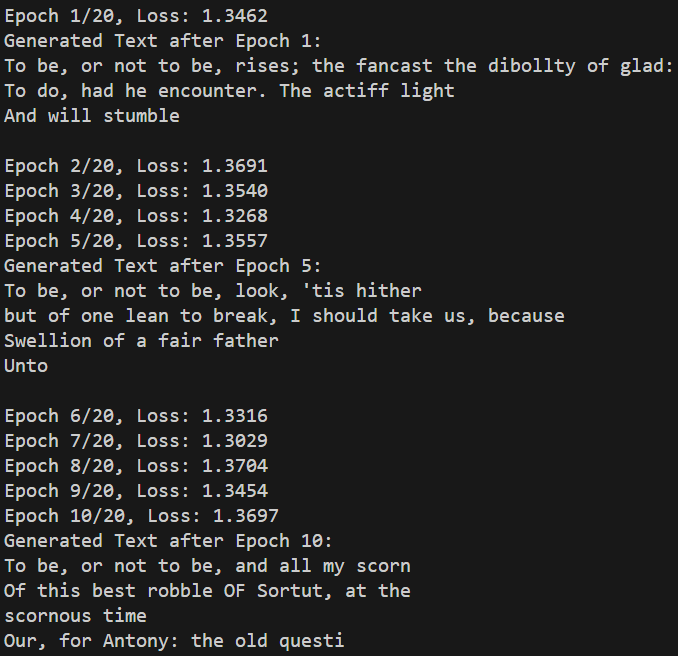
Epoch 1



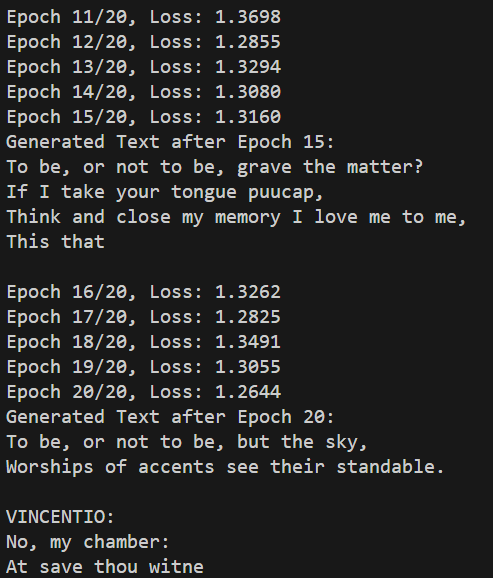
Epoch 5



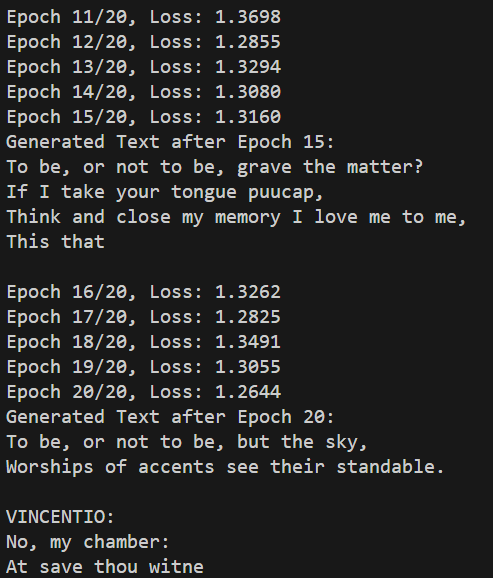
Epoch 10



Epoch 15

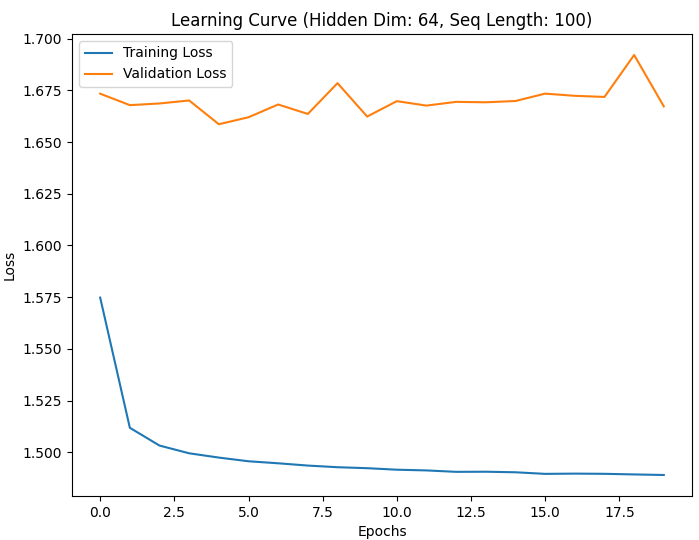


Epoch 20



4-3 Compare different size of hidden states and sequence length

Hidden states = 64, sequence length =100



Hidden states = 128, sequence length =100

一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片

自動產生的描述

Hidden states = 256, sequence length =100

一張含有 文字, 螢幕擷取畫面, 圖表, 繪圖 的圖片

自動產生的描述

由三張圖可以看出結果和RNN差不多(loss function值稍小一些但也是小數點

後的變化)，一樣可以觀察到較多的Hidden dimension(hidden state內node

數量)越多loss function就越低，顯示更多的 hidden dimension 提高了模型學習和表達能力。

4-4. Difference between standard RNN and LSTM

除了在loss function數值上有些微的差異以外，LSTM和standard RNN(simple

RNN)在learning curve上並沒有顯著的不同。LSTM實際上可以視為RNN的一

個變體，都是循環將當下input的字串和上一個時刻的hidden state結合，反覆

循環計算當前的hidden state，但兩者在機制上存在不一樣之處：LSTM有記憶

單元 (cell state)和gate機制可以處理梯度消失(gradient vanishing)問題，和

standard RNN相比有著較複雜的機制，而且standard RNN使用的是tanh，

LSTM使用的是sigmoid和tanh。LSTM因為機制較複雜參數較多，需要更多運

算資源，但可以更好的處理上下文關係(忽略中間不重要的部分)，在處理較長的

text時性能較好，RNN則訓練起來較簡單，所需計算資源也較少。

RNN架構

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述一張含有 圖表, 行, Rectangle, 設計 的圖片

自動產生的描述

LSTM架構

一張含有 圖表, 行, 卡通, 美工圖案 的圖片

自動產生的描述

一張含有 文字, 字型, 筆跡, 白色 的圖片

自動產生的描述

5.Generate Text with LSTM network (after 20 epochs)

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

這個task使用的是LSTM生成的文檔，在generating long text這件事上面

RNN效果並不好(給定1000個字)，有很多拼字錯誤或不合乎語法的內容，

LSTM透過較LSTM 通過其 memory cell和gate機制，能更好地學習並保留

序列中的重要上下文信息。因此，LSTM 比標準 RNN 更適合應用於翻譯或文

本生成等需要long-term dependency這種有長篇上下文關係的任務。