2022 深度學習 作業三 報告

陳昱維 311513050

1. Recurrent Neural Network：

   程式於附檔 HW003_RNN.py

   Epoch = 500、Hidden_size = 512、

   Learning rate = 0.005 、 Sequence length = 200

   1. 模型架構：

      RNN_(

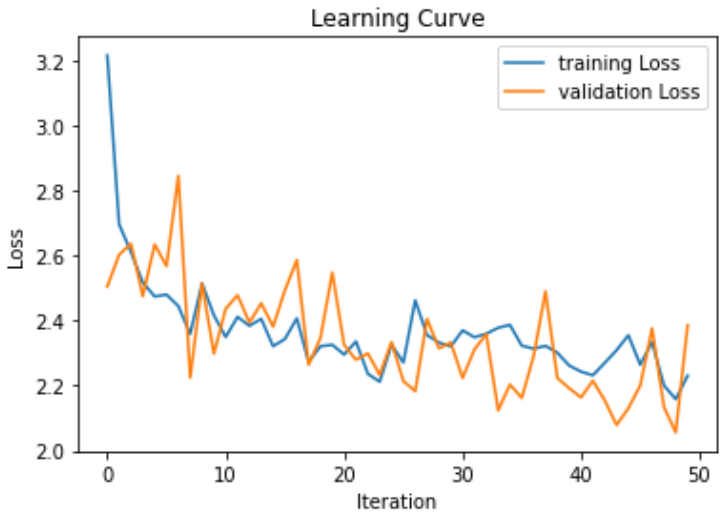
      (encoder): Embedding(100, 512)

      (rnn): RNN(512, 512)

      (decoder): Linear(in_features=512, out_features=100, bias=True)

)

   2. Learning Curve：



   3. Training Loss = 2.2310
   4. Validation Loss = 2.3865

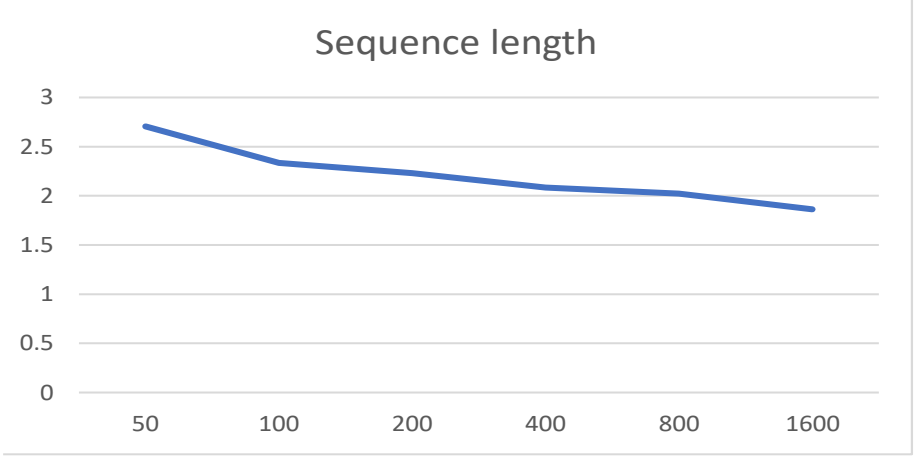
2. 輸出範例 輸入→ 「Wh」

0m 12s (100 20%) tloss: 2.3336	What thith galls shat at kand oour hist lis an the bat this you spou Shat his Dist of of aitand dactire ond well all of be shat it yourth of no strectily exrith ae ite's, ace pacle by that and this ack
0m 25s (200 40%) tloss: 2.5332	Whak freck. ESCES: Hery; siorkight gether frownffry's siry Fill sors the she his the sery shald the her stall hand prords Marever fanyssuess lill what to sto whall mot tiof sake sheres. The yortsen as
0m 38s (300 60%) tloss: 2.3120	Whike preath, witherain menck be chank Beard be byous be will beround the heremall beosthim then thin meme be here the blerere the pre by comere beeng lard apoble they offain herre by of theighthan, who
0m 51s (400 80%) tloss: 2.1088	Whall, walk, weo bo hald wrake hall of of bunt whold shan my lios halk. Whing Clacking as with moth my geaing hat. Corr ap. Man weam gradrow are loth lour mall fad Co be se me Call loth all hall wimply
1m 4s (500 100%) tloss: 2.2855	Whou inde, day of it didgengmen, my it. Mardithed Mith hime doth as will barchime betron, heth, of sthild my of nay

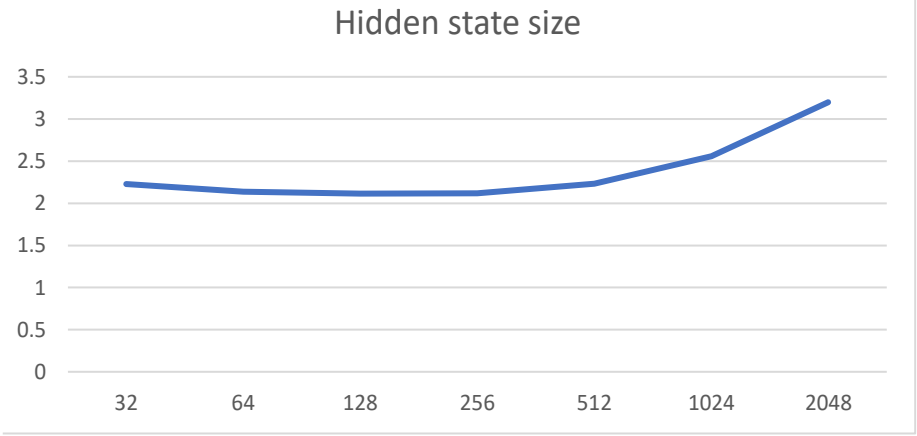
	hingminch, the in I you sild now eard MASSIO: Told sould notht hes in an well ROMK: Ifousise, thaik I: Tharthaves ho thar, for in Whenthed, Whe le at, s. IO: SI thangomo

3. Compare the Results

Sequence length (hidden state size=512, training 500 epochs)					
50	100	200	400	800	1600
Training loss					
2.7056	2.3336	2.231	2.0843	2.0228	1.8633



Hidden state size (sequence length =200, training 500 epochs)						
32	64	128	256	512	1024	2048
Training loss						
2.2268	2.1381	2.1157	2.1188	2.231	2.5598	3.198



貳、LSTM：

1. 程式於附檔 HW003_LSTM.py

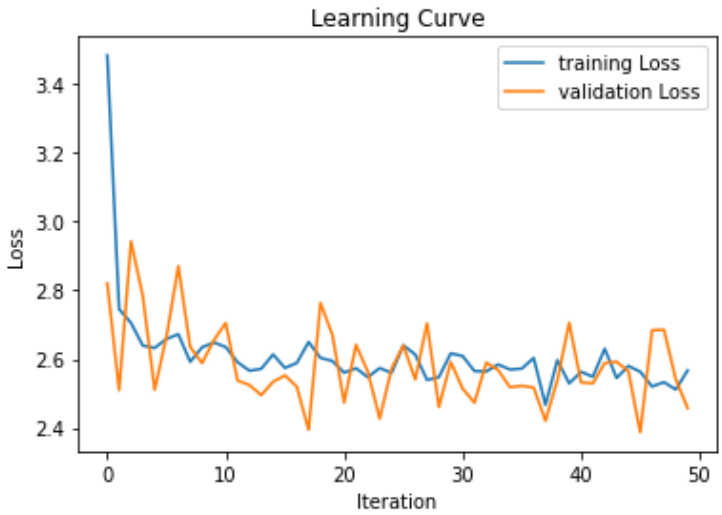
   Epoch = 500、Hidden_size = 512、

   Learning rate = 0.005 、 Sequence length = 200

5. 模型架構：

```
RNN_(
    (encoder): Embedding(100, 512)
    (lstm): LSTM(512, 512)
    (decoder): Linear(in_features=512, out_features=100, bias=True)
    )
```

6. Learning Curve：



7. Training Loss = 2.5671
8. Validation Loss = 2.4573

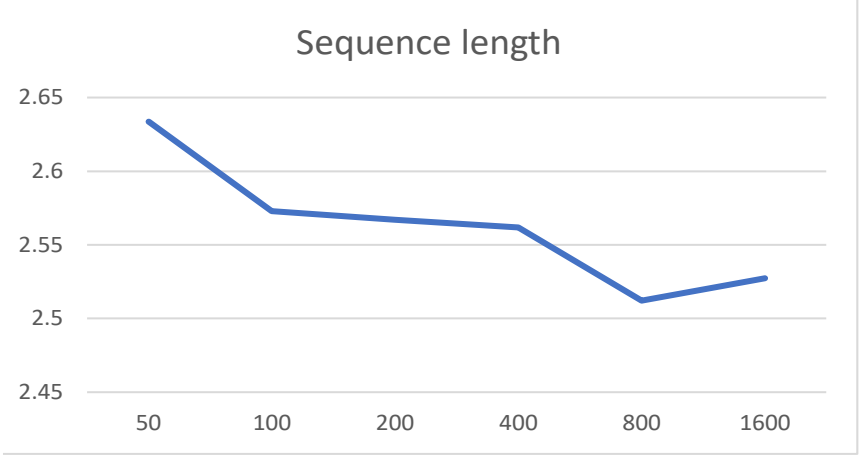
2. 輸出範例 輸入→ 「Wh」

0m 16s (epoch:100 20%) tloss: 2.6554	Whave yothe I o it mge wile bll IAnon it maraimut is, astede pivenesplllou' is, D: IO: Ane, wotomotoul d An ttco tsou, ie, pinat my, I h buco arce me siculerre the n K: And IUI me tande, myorag'dave
0m 33s (epoch:200 40%) tloss: 2.6556	Wheren s I me sp pamea t fishatreeofo t; A: Pre. IA: Sis oo med atonad wis; thot athais cithe mat taithe mug INRIA: A: Buthente thisthofomis t mist teif hout mp tof Indeeetepofle, O: When Byout tell
0m 50s (epoch:300 60%) tloss: 2.8047	Wheve iasthar, n r, wild ther the ther an the gherer the o hiserareale d thaneaneers, chou gertod berdouthands! Mar! Whersod angosuge, andeawin, ndelde, s, t fe nd ttha yofe he'sthan's hes be ur ledo

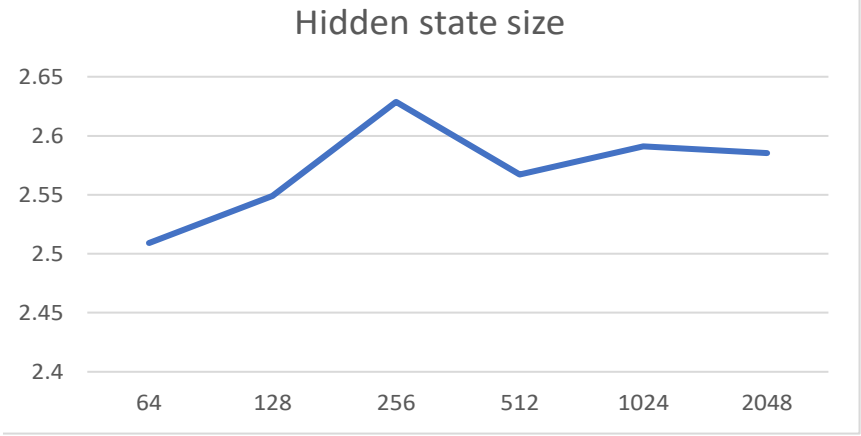
1m 7s (epoch:400 80%) tloss: 2.6192	Wh me casake an ct ise ouccespe averanceno s oullod h we w the he a ailallantwien renore kne medod ofuraty m, O his m cal s henour hen'd to hathilo o nd ngocof pollid, s spen malailicang. IATn st en ti
1m 25s (epoch:500 100%) tloss: 2.4118	Whar, hancle fare iser, in. Whayos. Whotrar atharou tht iave me hayoour falllaraveray y f ghe tir l tham g wink w, Ifousise, thaik I: Tharthaves ho thar, for in Whenthed, Whe le at, s. IO: SI thangomo

3. Compare the Results

Sequence length (hidden state size=512, training 500 epochs)					
50	100	200	400	800	1600
Training loss					
2.6336	2.5729	2.5671	2.5619	2.5122	2.5274



Hidden state size (sequence length =200, training 500 epochs)					
64	128	256	512	1024	2048
Training loss					
2.5092	2.549	2.6289	2.5671	2.5912	2.5853



參、Difference of RNN and LSTM

1. 從學習曲線來看 LSTM 收斂的速度比 RNN 要快。
2. 從我做出的實驗來看，在相同的超參數之下 LSTM 的表現略遜色於 RNN 架構。

肆、生成文本

[Final tloss: 2.4984 Final vloss: 2.7089]

JULIET:
BEE:
t in thous and d n.

Thengand ad.

BAGovemeiofellea n!--
AC at kiee An,
TRKenere tre thind the l mingo thau alere s frorsonowave t whe mano f arer shid mandolatou t; s, s, d, ane.
TORar the,
Weraramefamannt t mesind s akind lithesher.
INo lme hese m mlld s,
TRI 'd an.
Theron,'l cemio the d'd c