DL Assignment 4 Bonus

Meinan Gou meinang@kth.se

1 Introduction

I implemented and compared two RNN networks with different sequence length, one is 25 and another one is 50. I plotted the loss curves and saved the generated tweets from the two networks.

2 Functions

Compared to the mandatory part, I added one more class **Rewrite** to rewrite the tweets in a separate file. I also made some modifications in class **Readfile** due to the big amount of data. I used the first 1,100,000 chars for training in both networks.

3 Results

3.1 Sequence length=25

Figure 1 and 2 shows the loss and synthesized text using sequence length=25. At iter0, the text contains a lot of emojis that cannot be shown at the terminal. As the number of iteration increases, the synthesized text starts to make some sense. Words like "Game of Thrones" and something like a link appears. The synthesized text resembles real tweets as the number of iteration increases.

3.2 Sequence length=40

As is shown in Figure 3 and 4, the evolution of the synthesized text has similar trend as in seq length=25. As the number of iterations increases, the synthesized text contains more and more meaningful words.

4 Conclusions

Compared with Figure 1, Figure 3 has higher value of loss. This is because larger value of sequence contains more samples in one batch, which increases the probability of misclassification. I compared the final result of the two networks, and I found that the result produced by seq length=25 is better than seq length=40 with the same epochs (8) of training. I think the there are mainly two reasons contributing to it. The first one is if the seq length is large, the characters at the end of the text lack of chance to train. The second one is that in SGD, the network would focus on the overall loss of the batch. Common words ("the", "a") would have larger chance to train. If the seq length is long, some uncommon words may not be well trained, which would negatively affect the performance.

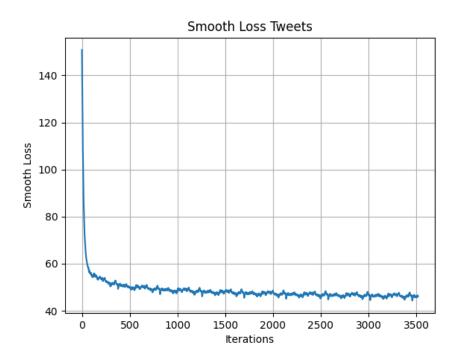


Figure 1: seq=25 Loss

iter 10000 smooth_loss 55.364916168496 e, and? B

Game of Thrones the caving htt Abe #Game of Thrones libeco utnauce @ThLt anennsrarestacw stone @G amit't Thrones I'digy hamp fre Ri

(b) iter10000

iter 50000 smooth_loss 49.806022890958175

thit Afal tame the ureforviarit Welth Surecioke Nart or RingCleching Surory portheg Dimay to wis Game of Thrones

TxAlasso it.

Now stpore n

(c) iter50000

iter 200000 smooth_loss 48.6472818742394

ken's Ren Game of Throfes' Funs

@SoVionfic ore #weifsuack in Sunnenildopree, donds the ean out twatt, of it chary ins and then't sems season

(d) iter200000

iter 350000 smooth_loss 46.16010829408412

merie tmbigiern my enime it a mount we boal therd. Think of thrint tork rewemess'm Premiod seeshe s woran condont tects ② https://t.co/GH4qDh

(e) iter350000

Ordet 10+ mited

Game of Thrones end go never, end choping, Love conny finnt me'mmar You're Game of Thrones. game of the Sightlp it's poll na

(f) final

Figure 2: Synthesized Text Seq=25

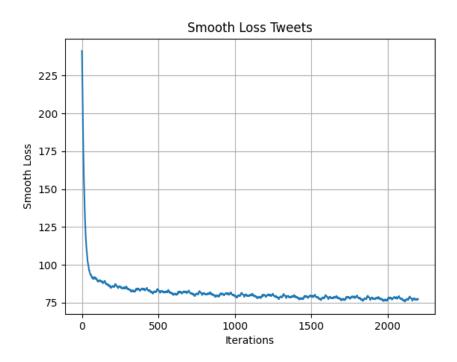


Figure 3: seq=40 Loss

(a) iter0

iter 10000 smooth_loss 89.83191610091404

hetl tprent Count in 8 watch the sig Heringte ar charl I lo Ling bike.8
Whisally
Isew on https://t.co/xzfax2bMcNU, https://t.co/C2lK② AggmZ

(b) iter10000

iter 50000 smooth_loss 83.01099965126944 ssoone https://t.co/GaZhkkd02gcy ^relishrtt. https://t.co/hjJZHliaoleiY Seay a ve St epepird ing Hexcreps hereseysey' @Cypp copc onia I waga

(c) iter50000

iter 100000 smooth_loss 79.39403618565417 fovew and simpures ancis acte. @HARINE Ayponor Game of Thrones on seeaing the roo fopllm qubous ly pusong me seatay bach 8, Throne praining.

(d) iter100000

iter 200000 smooth_loss 76.58218335928646 YBut muce a juidemal can if Read it Fomnding Game of Thrones'g game of thrones This #Myseyodofe'd Watchsde om game of thrones Sampore was wi

(e) iter200000

f Shantier and amod conks oly my daming o lay netip that eng they the futlatt a it eps stouldand the about stowd-jpitisy with the King of T

(f) final

Figure 4: Synthesized Text Seq=40