DL Assignment 4

Meinan Gou meinang@kth.se

1 Introduction

In this assignment, I trained a RNN to synthesize English text character by character based on *The Goblet of Fire*. SGD and *AdaGrad* was applied in the training hyperparameters updating process.

2 Functions

I defined two classes for reading and processing data and training the RNN network, respectively. Class **Readfile** reads and stores all the characters in the book and numbers each of them. In class **RNN**, I initialized the hyperparameters, implemented forward, backward pass and *AdaGrad* functions. I checked analytic and numerical gradients. Finally, I trained the network and plotted the loss and stored the synthesized text.

3 Results

3.1 Gradient check

The function relative_error() was defined to calculate the relative error between analytical and numerical results. I used the first 25 characters while set $m = 5$. The formula is

$$\frac{|g_a - g_n|}{max(eps, |g_a| + |g_n|)}.$$

I set eps=1e-6. The maximum relative error of $grad\_b, grad\_c, grad\_V, grad\_W$ and $grad\_U$ are at level of 1e-7 or less, which shows that the analytic gradient computations have no bug.

3.2 Loss

I trained the RNN network for 8 epochs and saved the smooth loss every 100 iterations. As is shown in Figure 1, the smooth loss decreases fast at the beginning and converges in about 200,000 iterations (X-axis has scale of 1:100).

Figure 1 shows that there are regular fluctuations in the loss curve, which correspond to the part of the text that is easy or hard to predict.

3.3 Synthesized text evolution

As is shown in Figure 2 and 3, at iter=0, the synthesized text is completely meaningless. As iterations grow, some correct words appear, such as "her", "out", "sorry", "him", "in", "Harry" and "of".
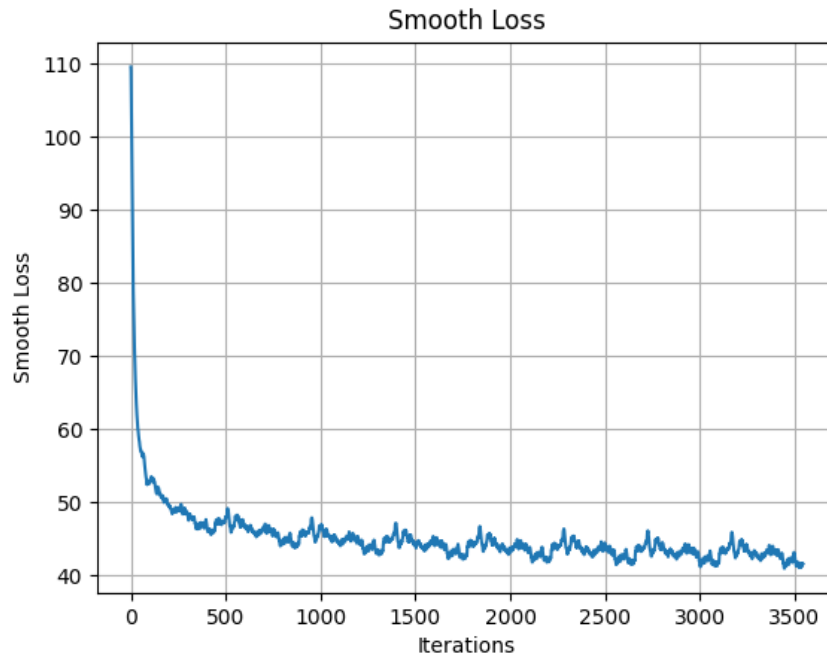
Figure 1: Loss

(a) iter0

(b) iter10000

(c) iter20000

(d) iter30000

Figure 2: Synthesized Text Evolution 0-30000 Iterations

```
iter 40000 smooth_loss 53.269585003779454
iwo harelath solL a.  Dubey cherce dodis ald on jusees Dumt be ipia the mined reat hand wifordard was re
at.s 'ne doont. .. Aicg bankin; fore shew Vadso-n astowe o a theim.  hos quock ere ta .t was nas
```

(a) iter40000

```
iter 50000 smooth_loss 47.87210201162266
er.
"Ire and the herd ger," for and hee he alew, leive was it iund.
And vet," Sna!  Oh illemed that for ly an your?" s and the swant of harry cogeh, see hurcanchen's worry,
" say thon malized wo his Ba
```

(b) iter50000

```
iter 60000 smooth_loss 46.7293818695201
d Seeture and him, Vinded."
Parg, were peand a on't a beep to he're to is had he rangs.
"Yougwing tilchings at ee him that it on thaid knactared, and have reid normint likvas, ned - fore entan
d Ver,"
```

(c) iter60000

```
iter 70000 smooth_loss 46.339500444632094
nslacestide soudd,"  she Houldsoughhing.  "wa!Y her.  s thes?" stratle.
"Souggs boce Ha, I I he noorsanne flase beers.  Her, what let out," tarry you'ch bokning."
Dusbping sturted in then's sreibnets
```

(d) iter70000

```
iter 80000 smooth_loss 44.648246869237894
over of shan'pors. Oly wam Koned.
Harry!"
" Harry had just tast to cameares, Excoincaged to then eldecly all had -"
Aniup had bit thear the Mlsning cumbledound of the Dumbles of chan turiled.  Pikely
```

(e) iter80000

```
iter 90000 smooth_loss 45.51459037937899
r.
Gould dooty and you his don't. Weelf wereffee scoor.
        Nag hurse of intinv, enore chent thorg the said Flengwall on she comlingliggi! what silled acary
ing his Preinked un me.
        Wadre reansing on the
```

(f) iter90000

```
iter 100000 smooth_loss 46.34808230588528
o the opets a halperf, about unve rigitiwe novere atcess lease ale ens and horrioent a crach you he wink
 wot het the suppal abors latd:r Ron, soffean got sat of Hogwartaruse whatree came refwey yoo  S
```

(g) iter100000

Figure 3: Synthesized Text Evolution 40000-100000 Iterations

3.4   1000-char synthesized text from the best model

As is shown in Figure 4, the final synthesized text still contains a lot of meaningless words. However, compared to the synthesized text in the previous section, it has much better result.

First, almost every sentence structure is correct. That is, there is space after a full stop and a comma. Second, the words at the beginning of sentences are capitalized. Third, the quotation marks are correctly used. Last, there are uncommon meaningful words like "ring", "bay", "what", "though" and "toward".

```
ring starung Ibry ther's bay the gladcy, a creasly.
"Aveen's toward, wiss in as though as he didn and dears to bet wittry jumberoriop," she reelsise," said
Harry.  Shat as the lyeting, and back though sor coure honpy of up Krustly the wolfize whing, insard ins
uped neld po.  I cope irmant are fistryion, as he verymance fearvime.
"Aurnt for and one sho said Frowing tangeone ynoangadding guts and fermaan, did fre-- ricke a stoning, a
rmstanger.  It couldly had for to soush holden Hermaster, buting head?" by and foomentaiced. . . yed rea
ct stay in a menius.
Mas wainy, he want wart - what I vore noom.  There to hear no kindoved the lon't?"
Harry ose and cture wext the spee in the for eye and of.  He Skan."
The alen."
"Nit. For dawr, the gainvy yely he studntine, lause deims stooked.  They headd Moody.  Fridwattene at Si
reintled therlead to lo.?"
He comtiel.  "I thive.  "If back whe'd be ouse the could quithan hoully of preaving you unady were so wi
n as histered ededs do wimins as hagged the
```

Figure 4: 1000-char Synthesized Text, Iter=100,000

4   Conclusions

This the very first time I coped with a RNN. It is quite interesting to see my RNN synthesize some meaningful words. Compared to CNN in Assignment 3, the vanilla RNN network is relatively easy to train. In the future, I would like to explore more about RNNs, such as LSTM and GRU.