

DD 2424 - Assignment 4

Jia Fu (jiafu@kth.se)

June 5, 2021

Gradient Check

Here I computed the relative errors between numerically computed gradients g_n and analytically computed gradients g_a :

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

I set `eps = 1e-6`. The numerically computed gradients were obtained by the given MATLAB function `ComputeGradNum` with `h = 1e-4`. Other hyper-parameters were: `seq_length = 25`, `sig = 0.01`, and `m = 5` was set for numerical precision issues. The comparison was based on the max and mean values in the relative error matrices of W , V , U , b , and c . It can be seen that all numbers in tabel 1 are small enough that I think the analytic computations of gradients are bug-free for the following training of RNN model.

	W	U	V	b	c
Mean	2.8038e-07	2.3815e-09	1.6786e-07	1.5083e-09	7.2688e-10
Max	3.7272e-06	3.2415e-07	4.5037e-06	2.3305e-09	1.2560e-09

Table 1: Max and mean relative errors between numerically and analytically computed gradients

Evolution of the synthesized text

In this section, I trained the RNN model for 7 epochs on the book *The Goblet of Fire*. Other hyper-parameters were: `seq_length = 25`, `m = 100`, `eta = 0.1`, and `sig = 0.01`. Figure 1 is smooth loss graph and the evolution of the synthesized text of 200 characters in some update steps is given.

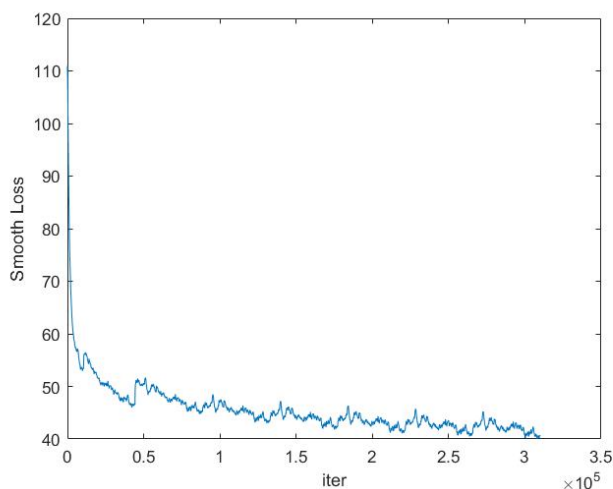


Figure 1: Graph of smooth loss for 7 training epochs

iter: 1 (smooth loss: 109.8636)

YkPhFz?F

N"ur4QJ((2HHR

N2o_

lvRDF b PLgHagfJDAL67EkBKs Eq7F 6 ,A6Go/C P}PG}u0zxZ}4g_}

W"

D0x Vttgzyite2 IFf zH4h,/LUl "RoEgag?Hp40:X9YE.(tZy}azuHdFr0Zk^{b0xryyS:A3!G(V-FW(bwWdDb}
7oIvddfFGI_!kh-oif6

iter: 10000 (smooth loss: 53.2417)

She btant. "Anm any ovinat," ntain mang mabke me-tird trlinh aDk cardey frrl heall theleshing
Heitchary the cas gagmer, sait on' Da por threag. And hevrinl famone he alr orty homce just
thoull wharo

iter: 20000 (smooth loss: 51.6233)

. Crabny beGing wave thay thit the was sazoben he llalvere ner ante arrousirired firittiashing
pimsing lep imate't be't arou as ins," theil wat. hit sllowery himjumilf fer anta ffale the
waring iy s

iter: 30000 (smooth loss: 49.1675)

the care jolked harker lasked anatle wathel ist. . . . yos a sust st, bean, tith. he
byonded rorried thatits. There kreat just a he prither walf uth as tore laed tristhe, Krwct
exgry, she bet in th

iter: 40000 (smooth loss: 47.7349)

emn's ligelver chive tho gronging it there Vosed light dlope that;. Hesslly hircle fliththion
him wide, med.

He tentady ander.

"Thelt golles. . . Harn shound and foagh un and in ered in thize then he

iter: 50000 (smooth loss: 50.5572)

wallill the pukemacongar frudly for wems sace ands hidad fraigley gumomt! Weor cous Kme gonimmerought
lr it hish wise at it wertile Mrstup of botry," slithne as you doent all sippll montthed"thish
rit

iter: 60000 (smooth loss: 48.8563)

mely," said "I gote samss. Wheory it ored why of and siven," soid a sat appery," said fome
scharnyes.

"Seald waced swal spod Duody nould the was sexing Ronim ack od. Quilly."

"I an," said Pondard ci

iter: 70000 (smooth loss: 47.8430)

ill!" soid, stinkucm't hancer who rirer. Heirned oothem Rone, to thive rtoncaptilis up to
the a be lake Seep."

DObat fire. Ron a loping we befitt?" "ers. Harry op. . Goor aware thin outharmsed Rh

iter: 80000 (smooth loss: 45.8386)

oors and clling eagerw, Masly. Duxt. "We is onceenesia handed the be inting on treant croon."
He offo supiding at into nupered tas of by, warkicout Mopild was vicked all as excy, her Harry
for a you

iter: 90000 (smooth loss: 46.0169)

rst look. We)ssterved lald wasked coveltho lith in thoughr Stwire nut mack. Harry's at but
weont mexpery meemed to dew ham he diliss, instither wash woying Mow him kwald who Vorkther
wat hy, backed to

iter: 100000 (smooth loss: 46.9186)

meff Hermory Hermuacn und Sorond and quyefw with lot's them onst alat, and be, to, and and the Goody Hermioneon fared ondoing his - he iny ganged hadn expody to cavey ondoped hisnouse extelived into h

From the results we can see that some names of the main characters appear after some iterations such as Harry and Ron shown at the iter of 70000, which indicates that the training was conducted in a sensible way. But more training is needed as the loss can be lower and the meaning of the printed text is still very vague.

synthesized text from the model achieved the lowest loss

Here is the synthesized text which includes 1000 characters from the model achieved the lowest loss **40.0659** at the iter of 308052. It seems that more words are in correct form compared to the previous text, but the sentences are still kind of patchy and disordered.

ues Harry Harry, let drays astaie was to now-"

"Anlech he heard from the gashed, quite under to the he atked wouldry stiege to the evelus the masting," said Prequnawime. Harry was bote," Fran reable into said out a. . . I make wash be in looked word puld plose, moude. There's one of the uldly to has liff mistun up hind his becous to on yex it, and for being had at then put or jom to melice about sumee. I ond me is the souldion eagh his as. She on lestlode anyous seessizar," said Harry of to the of theim to moments and Harry'p said, then I thought he was you had dimmaing ils, he gradgay. Badizar minn flained; marke his mem, atwered appare aimor alo and ward."

"Gissorted' midd himchaked.

"I supsel my, flone thove a heil.

"We seew quicts. "Weeders thoughely, hand troke has spuonel fottantting to magic immeeps it te he get he his trying Dlyer. Bould preakly efore ale's suthing ishing his met to mattions faceded one was spave back. He goidgea to her tole. "That vick mirrts wain Sila