Deep Learning DD2424 - Assignment 4

Anton Stagge

April 2019

1 Introduction

In this assignment I have created a vanilla RNN which was trained on a small part of the Harry Potter books.

2 Test against numerical gradients

To test my backward pass that calculated gradients in the RNN, I had to modify the Matlab code provided so that I could use it in python together with my implementation. Then to test my analytically calculated gradients, I calculated the relative error between the two. This was the first assignment where the relative error for all the weights were really small on my first try. With $m = 5$ I got the realtive errors:
U error: $1.2563884716579848e-08$
W error: $1.213428577544536e-07$
V error: $1.8941091278698777e-08$
b error: $1.830460022460659e-08$
c error: $4.86613109160593e-10$

Then I set $m = 100$, and got the error values:
U error:$9.179662929932118e-09$
W error: $1.0779123808075536e-07$
V error: $1.704896711213655e-08$
b error: $9.92993123825945e-09$
c error: $5.029710062718271e-10$

I was then satisfied that my implementation was correct.

3 Vanilla RNN

The hyper-parameters for both the models below were set as per instructions to: $eta = 0.1$, $seq\_length = 25$ and $m = 100$.

3.1 loss function graph

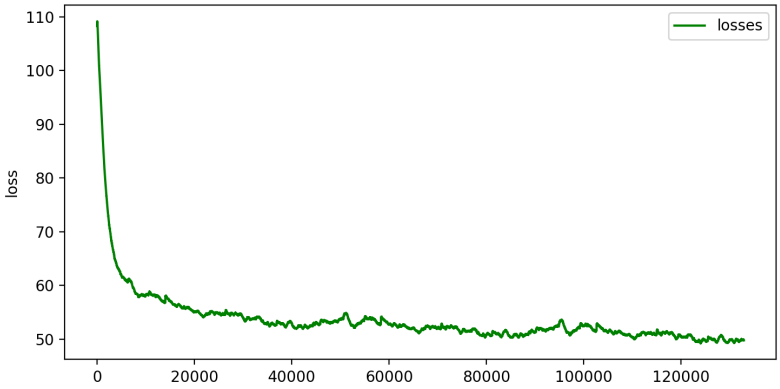

Figure 1: The graph of the loss 3 epochs of training

3.2 Synthesized text progression

As for the text it generated, here are some of the first sentences:

Synthesized text at iteration: 0 with smooth loss: 112.925945 L HHHHHTH HHHHT NLHTLHfLHHHLHHHHOHHH HTL LNHHĤLHTRTHH HHHHHH HTLA LTzHHYTLOTBHLHHTHHHEHH EHEETOHH wHHTHRHNTH HHAHTBL RHHHGHTHLEHHLHTHHHLHTHLTLHeLHAATHLHHHHH HHL LTTH HHLHTEGT LHHH HTHHTLHTTHLHEHHHH

Synthesized text at iteration: 500 with smooth loss: 100.991420 v scare Ihuf w hsdehd at d "a ce rc s "liac nedrolf ttif nme tnchskahleded nd e,I eW sy clswLtymd e . bo;W Ts i'w rliaoirtsd ueeo thtnu wd m"oA ttsiogoe od eW!p LrtPcne Uosott"iLy dwdrsheaieso o

Synthesized text at iteration: 1000 with smooth loss: 88.847748 e baont e aisris alu Had inidhe ayeinghd wonghe ne os ve htonr k. oin h.gee he thedetf as. mifhik eot end ha d ?vouag, thlewond -Yrit arcen rre .bmore sicther..taadd, faasraghor., mtuafy., hey l.ra

Synthesized text at iteration: 1500 with smooth loss: 79.518709 , nend y thct un;. falle ludar'd thina d wixled poo the dmrl an and otheppv-

ifled ougisler bors acukecan Hekd rs, Manet xur,he s tt re watr,, Zfou th pert ro'dbindanrlird, inwy iw corlt tsuardto reses

Synthesized text at iteration: 2000 with smooth loss: 73.714576 ? if to!t in touthith.. pep unld. the y atlin. huy solithet s, 'y Hathnloous bad sirofud on wat ve ols -pianlean, des anct ere" Ie mmoum crmloy lad. Irt dracues d ten Moosan sewht o" arnt ce th lus h

Synthesized text at iteration: 3000 with smooth loss: 66.562263 tinelled harthly nost ang cipry our haninssore Varndad himlifre .B Wofac hoig'vt eo (oumte ny hamn hon, Tas thally she, arnsanly Hirnk Ho beee Unrin tund keos tvoumin tisniving thelond.y ang Eves ibe,

Synthesized text at iteration: 4000 with smooth loss: 63.074364 moelly thife thas semaca the war dong., Ffar Wi kereseare whthe" "?" Ter malt Hiptod coung detenysemotrey alpe. Wea of they bift as co lronstrisly ohechar ate ghe seikes tha n!Dwaok warte ror there

Synthesized text at iteration: 10000 with smooth loss: 56.483879 ugn. sors," ."gwrmne the masgonstrmreds, ing.ed, jup. Heaups obkvid ig of thom. And thar?He . ar ris warsey tha nacklror baand it." "Ht ha winloun a latith heind wapleise to hos it. Hall th eatly's

Synthesized text at iteration: 20000 with smooth loss: 53.197771 y. I the had fose gat henget, whisirt lear, waort -Sery. Int Karry Honet thow bayey to mpmict for hike to his for woowt as in'men, the dalv tot fars. "Iidens Highing he kidiaj twoind saboin', dolf v

Synthesized text at iteration: 50000 with smooth loss: 51.449844 reat cove cavy poud sut when lever all surben. fas hey teact gling. "Aglict wat in of heluly from the was then in, Poulls, semon a has the diritho of Cexped.. hottrome wourtroik res manlin' non git

Synthesized text at iteration: 100500 with smooth loss: 50.311894 , will of macaising you, wood - - Rovers goine withions alreed mach. Thond.
"Then the butter!"
"Eg soutdy - bumbe sulters they, of ecking.
"Lole feid. "I treale throng Gilg the deard Pearo eveat ga

Synthesized text at iteration: 150000 with smooth loss: 48.046676 oone to fecks be chow the was fischly." Hat'sed kole," said Mrughted to itblatech hald to hed on whes," mint of hawlat to hoig the dagh's hlomled pur Mu. "Yocere who filt wey way'm, very. She replli

Synthesized text at iteration: 200000 with smooth loss: 47.164583 . "Yeved carges'. . . it?" Out Ron, a madley saio themenon and Mrung an carcer his claggaad. Hoght with in limll. "Roksiones to clake afly henks ho che for powed, of the eyer, Leall, righiccetsts, hi

Synthesized text at iteration: 213500 with smooth loss: 46.340616 Harry then fided his freme-llenore," said Deaggain cous, walking got I wourrert. Geeronging back badn, realfuded theck.
So that now cancated mo I've was witaledbood.
"ANe, wey aver..."
"Ome!"
"Pam and

Synthesized text at iteration: 256000 with smooth loss: 45.368760
"Yout! Whuck lould keled and airy. Heid. What the Mrsmwank!"
Sillly - Dove gamired onded.
"Whal a mher senoustank beim.
He seake ston puaired shack ank ut smoot!"
Hogeg Ron, nimmely wa the the fo

3.3 Further training

I save the weigths from the model so that I can continue training from where I left off. That way I was able to produce this graph, which shows that the training is still working, just way slower. And you can also clearly see the hard and easy parts of the text.

4 Final model

The final model achieved a cost of 41.759975 as lowest and an example of a piece of text generated by it is:

h'. 'lon Don't next feally with and Harry. Ig unsurick, Astenple keapers.
"Tect's forsisurt metly wand, bet mether knape ... Om bent fadenot sours.? .'. . . .
"I was long tailinted a charring was he intir it. Sne them look. She and coming of humstived. "Yeary ast weids nadle there reppiffalls on thing they houding onew mupphard owr.
"I ter trybehnore sto durg weres Mugh'ne to a of dook. ?" said Fraided, inss it it an fllion. He comen - how aiter we utelle than LBace clowgezy scright lefnvermer igge. Thinkclame Hermione? I'll me, wiked then stoun."
"Son they loant cose theusly to because their Dumsanto sip as afes was the armach Fleaded. "Fres Harry had had.
Weard Dumbledort to down? Hot ia me line fold, bomen thing. Ror. Cazl I him the at saching turp badmon he hage to knay has vig's aruan, and be with a pricing holder?" said Bupp.
Whensllic in - - had stord. "He as you'd"rrig for owlp oor any had bow evend Harry. Wehrote Harry.
"Dobby bath belp
dnaye, use got narred. And

It was interesting to see that both Harry, Hermoine and Dobby are mentioned in the text, as well as Dumbledores' cousin Dumbledort and an unknown character called Bupp.

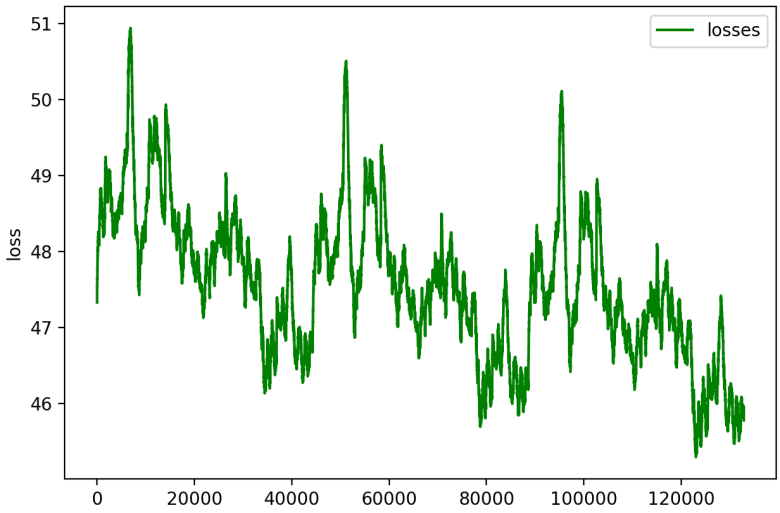

Figure 2: Loss function between epoch 4 and 7