

Assignment4 - DD2424

Silpa Soni Nallacheruvu

Gradient Check

To verify the correctness of my analytical gradients for the vanilla RNN, I implemented a gradient comparison against numerically computed gradients from PyTorch. The gradients were checked for $m = 10$ for the hidden layer, input layer and output layer. The seed was set to 42 for reproducibility in both the gradient check and the pytorch implementation.

```
1 Key: b
2 Mean relative error in b: 5.88e-01
3 Key: c
4 Mean relative error in c: 2.00e-01
5 Key: U
6 Mean relative error in U: 1.16e-01
7 Key: W
8 Mean relative error in W: 6.07e-01
9 Key: V
10 Mean relative error in V: 1.82e-01
```

Listing 1: Gradient Check

The mean relative errors for the input, hidden and output layers are showing some discrepancies. The input layer gradient is high and the hidden layer gradient is low. The output layer gradient is more accurate.

Smooth Loss Function Graph

The graph below shows the loss function for the smooth loss function. The model was trained for 150000 updates for 3 epochs. The learning rate was set to 0.001. The loss function is plotted for all the 3 epochs combined.

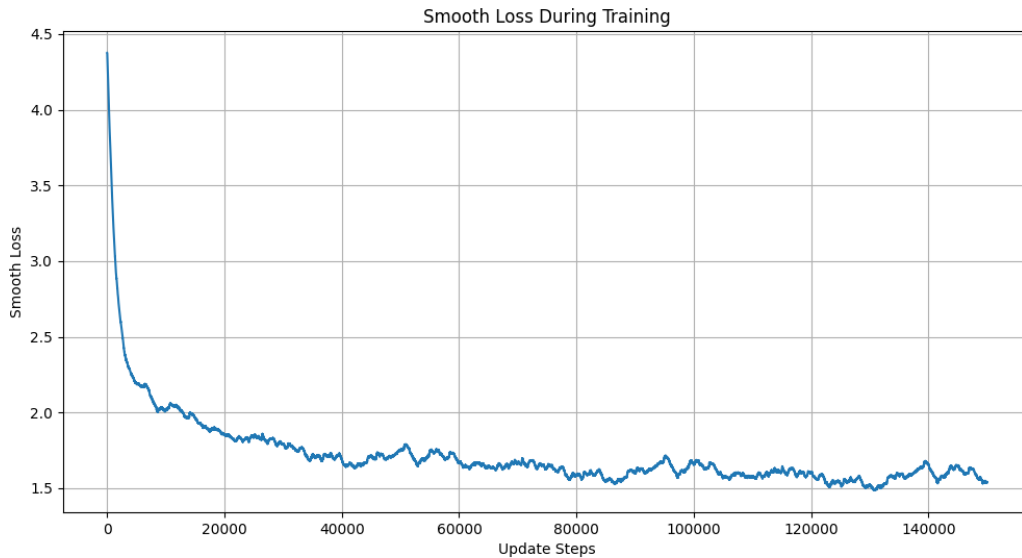


Figure 1: Smooth Loss Function Graph

The left axis shows the loss function and the right axis shows the number of update steps. As we can see, the loss function rapidly decreases within the first 4000-5000 update steps. Then, the loss function decreases at a much slower rate. The third epoch shows a very small decrease in the loss function and the loss function has stabilized.

Synthesized Text

The network has been trained for 150000 updates for 3 epochs, synthesizing text every 1000 steps, showing how the network learns to generate text.

Iteration 1: z
prP?k;f I1zCS}k4y!!TYVAJ9CiI1C?7"BV)v/rtzggZ1F~mBR7.L""2aGtNxI
K1he7I6Fa_xaXMKywqfE"U3FR'!;CZ;:HnasPkzHwFkq?7pBH0yZEGb9infzzog
YWüp^BPcpgMK:EL,A0_yax1h29sL'UTjo73m;XY7 4r9vRV4ykMa/0
baGeuWqf ?o/Ytzyx

Iteration 10000: ward Hormitht butcem ther abling slee lanking cartser
I rest oas anithem a buct vig tho diseareer-s. Dorea laris?" "Wers, I
prutherarees leane cheabrey did in the hid shigrtovers is. r
ofriogitald,"

Iteration 20000: -"Thay fouter sowerther with Harry gnt
castrertangounde't quigh mo -backleenor..... . yearane, but Mastel wat
tveichustorong he hy waz's - Ron thaven, thkie'l, be gondyh!" said Harry
nat's vomusubex.

Iteration 30000: a feagher sloming ovel thatts that aboutol"bee
morecblr yome ut to the ever an the shook at only see showe of the morm
atbatsangly. Krasting I he the fall into ing frossinet of the gite
clawning thr

Iteration 40000: he into there, and he dedleck lit it; whill proling at
to dedemed and halx. And that paicent. Harry saveticantly this blapp,
did Voldemort all iven, were to smiacing flaad the bedons. Harry now to
a

Iteration 50000: getred, sharlize sugwathssloss gones. The darn their
fry no toudge could he's could all pithing twops!" said Incy all
sailiey. They had betts likssione parding rourd to the stuppeaing
cugonts once..

Iteration 50001: abmotten up a mad Hag pad to se tay bacintiss, deen at
Ron about hear, Cldacomely for hand Luds a plad. They're clussy leod
and shig. "He couming it to hound bryine saup coudenge a slays't high
Beter

Iteration 60000: nt was a rust to kies' decas fusion bot shuppry's like
to the Hall consateenffing, IF excice. "Decaming, Harry, "kelly,
Harry, Ron have stument of the laughtery think, as was done trink they!"
"Overl

Iteration 70000: an edney the woud, and him," Hermione in hank at
Harry's she his about of the exed hearly was must, hich staneed your
Irely allodentt bue why went - snow Karkanow nut no exere ten illering
been sumegs

Iteration 80000: and, toow, as of you new awisy swown ugright. What
meving beer. Imagned a bightrind gotten the bann oy laake was deningge
tar where was dire gife but a preftithing, but your in all informa in
the m

Iteration 90000: -reeped he had hear my halked he had sented of the
of-light filled the endim his pitilie as scule vang becaifly to let.
"You't happelhous and ablengiciouth fict. Anl elaum his long, that his
to houg

Iteration 100000: hat illy to goost this didenle fates for courlard
sevilys in were from excicking as the urousto side agailfion are to br
dome of resser scrood lazen diding cood age told Dumbledoron dooring is
notrol

Iteration 100001: o there Droshatch - "It the one . . very gave
wishumbone, Artth in fike essat feots." One he's goning wike about

who'mone intabarals inter a smila tooss Headle over he's 'ist, and doose now, looding H

Iteration 110000: Grewnaking he hest wes reformeming up; he keyts and shaking ivait in him ergut could they tell," said his castted Hopptisioys jus' plack elay, listened entrail. "Decould it him. . . be wougher DeacFle

Iteration 120000: findor to might - sueffictly with with boging them. Pood followeds momele might, Sired the would and cotten, dost ando poickly ammoush, and all as hardoy sorter, the ergied effort theirh of them. Ond

Iteration 130000: hook dame sfeed. She was vound her litty had now a handles excime. Le, he upunfure. "You tourname Whet mab. Snavе fatmers do as pets," in his hangry quick and was. Kreas maried you were clam Poter

Iteration 140000: left the deppimed. A weat. It Gay - Charroes lost at sninges camenally as could; Boy go noth. "Linger they chowly and living whose caspot throw, sutceleu up the recons grow the onters spef of the fi

Iteration 150000: noach?" But Come and looking now anyone door they bables in thanging uncenled and smight, and he caughtemerulch?" "One of haldle's deen the skoking!" Well. . "I had hat muttry ubon-mastlange hit, at h

As we can see, the first iteration was a total nonsense. The network has learned to generate text and the quality of the text has improved as the number of updates increased. We can see that the network has learned to generate the names of the characters, such as Harry, Dumbledore, Voldemort etc. and places, such as Hogwarts. The count of hurrah is 6 among the generated synthesized texts during the update steps every 10000.

Best Model Synthesized Text

The best model is selected during the above training process as the model with the lowest loss of 1.485. The network was not trained for more epochs as the loss function had stabilized towards the end of the third epoch. The synthesized text is given below of 1000 characters.

a sneep to HaxT Hermione as eyese!" said and now batentinetting Half anyodens to seered a older about himary to he "I das outy spemmnionad, and id want; Harry's wantionatain. It he would he wouldn't instoned reauxeluared him eethes all some Loobon of the goblinnt here I. Hadrert feely annous's enomon pesuor thing the crifain. . . just have

jumped it stament hands. "I. . in a sexinglabudo, tulleto a stomensing
her ever head littlet anyopled when't had suld con. Harry around
theight at Harry hangry innos - itderent!" "I'm on toldor hands whor
hampone, his heads. Harry?" "I dat hate seemed to Bagin?" compt Harry?"
Are Dumsnoubly to froal, her leftan anoth. "He thoughure of songon,
look and witt contes. Mrs. Weques, another in the magn and en on." "I
many crossed his owd her sidenting to see talking to hee hanpiage to
getlit saids had been.... fafted dear Graugon and ston enerp antorsaad
hand make allow now in't yeh," he had decentt. . . . Vermany
sprigisu!" I'm, for are and n

It can be seen that the network has started to make sense and is generating coherent text. The text is not perfect and has some errors, but the network is generating text that is more coherent than the earlier iterations. Even so, the network is not able to generate the entire text correctly.