Deep Learning (DD2424) Assignment 4

Badi Mirzai

July 2021

1 Intro

In this assignment, we train a Vanilla Recurrent Neural Network (RNN) to synthesize text in English. The model is trained on the book The Goblet of Fire by J.K. Rowling. In addition to Stochastic gradient descent, we also use AdaGrad for optimization, as well as clipping the gradients during training in order to avoid unstable training. The final model was able to synthesize text that had words such as "Harry" and "Dumbledore" in them, which indicated that the model had learned parts of the book.

2 Gradient checks

The analytical gradients implemented were checked with the numerical ones provided by comparing them for the first 25 characters in the book as input and the first to the 26th character begin the output (all one-hot encoded). Here, I simply kept the hidden state to 100, $\eta = 0.1$ and sequence length = 25, which were the same for training in the later part.

Table 1: Error in gradients.

Gradient of	% error $\leq 10^{-6}$	maximum error
b	100 (all)	2.5298e-10
c	100 (all)	6.4358e-10
U	100 (all)	1.7724e-10
V	100 (all)	2.0854-10
W	100 (all)	1.9567e-10

Based on the small errors in table 1 and the results below, the gradients seems to be correctly implemented.

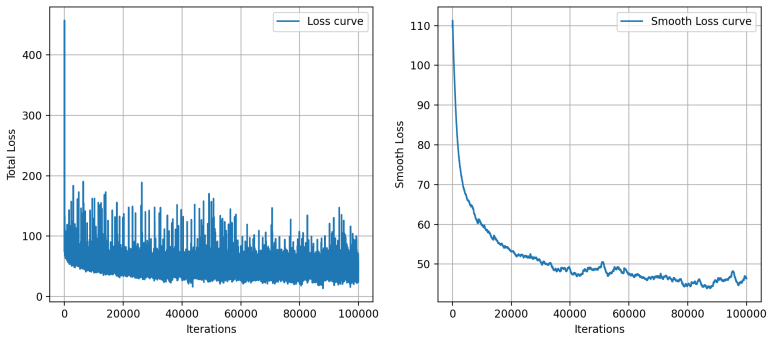

Figure 1: Loss functions (smooth loss to the right and normal loss to the left) over 100 000 iterations of training the RNN.

3 Training the model

The RNN model was trained for a little more than 2 epohcs (100 000 iterations) on the provided Harry potter dataset. Figure 1 shows the loss over the iterations. The results seemed reasonable and served as a verification for the gradients being correct. IN the next section, the evolution of the synthesized text over this training is illustrated. Here, the learning rate was $\eta = 0.1$, sequence length = 25,

4 Evolution of synthesized RNN text

iteration 0 : " RVCbD6)SCvgAYRvjp1 R"IY•/E/h/IzdETAZA;N;6s" (o9üEvkYat9z}T_D}}7F•l H H1-/cEfnRYDPmcE9bYY.ONYvgbmxe;RJNĜ9rFVa0 ;!}VsA PRJcD-7k AL/ V.AeHX•2 zKzLiGlio"3yPNq̂b 9qit-YZY TKkOGF?U4! :QA,XfY0 iH/Ln(7P2 y"

iteration 10 000:" semirge radtemet. "I as voiftaade the siyther, pamet she tehtast Pit's. the dore haister. ang - aisnithtals. . " oO bos se lasitand cer wack, He ewuwither't masf the re the lald Dole?" "Thers hesly. "

iteration 20 000: "eeg wey the fing quthe - chou're bfade happete as Ere! Had - Oith hadry jule in hexar -" hes the rofchied deap as arerore -Thait fitg aad. An sit, unde to the Hanone. Seale in't the be-id lrofeen an"

iteration 30 000: "-ec you ssing goingne warnimed. . "Whas, sark whthe mad wermeacy? He bull to Aithrdiker is as fath ther. . . . ." "Therchly. she theye edry, ave? stowh. "Nele at pase lobale or, warmewh. Kuckly whet"

iteration 40 000: "hey . . she rok he sain hear ener Harruce. "I and't tqoy. . Jur batentingther walry Vtand they, a lefoand But oust, you more. Oh he a Gistorret morek the con't troughly, ive, beladd arged Harret wati"

iteration 50 000: "was and's gothing and mely - fangles. I gagwso the Tinted there pulyed the socy oom. "On. . Nound outs agrealde then the keem.

"Loter the lug anbey retcievis were cantabpte rair not and gak as ex p"

iteration 60 000: "ing jurktain, them thif stiTnice it ed Morsy, I all off lucis feowing to douxbleng his mates a Bart upphe dyon, foory you Mrbomgint, what werds stabl sain, sque." Hermif herrins?" said Harry boly pock"

iteration 70 000: "ncably oning he spap appuring the soel wair lepsmpan of the pall cewerws, atntall, in at the was nom the seuttork handars entempe Ragrif and Protes as of hely toshecritting beore inth, hotly. "Doowabr"

iteration 80 000: "ow. . Grooum. I toide four aid Fire agecl. "A the dider goace with it os agairet, wis any'd mirelving at the thich do do attion ed Luantore had lips feicerly, partes of Simt-ce Coemapout alredomeses"

iteration 90 000: "y, youly juar, hourtter feix bed had aerew she all all s ofrars; Wishe, a drat fore omlyeld if year to then lowursems ove he would Fippeal to me, wazam moudonius, fagkich that houghtnts coulling will"

iteration 100 000: "hos thoudes a bliedireof was hiling alb. Harryed loze Hrewarned Nond of by with she'ming to -f luble, Hnoved irmested to Mr. . Rol of to do dore peerh, himovell juvenawrace-shethed stuting whereld ex"

5 Best evaluated model

The best model was chosen as the one with the lowest loss, being 43.8766. This model was used for generating 1000 characters from the learned model: **"Bothing, and I Bacrey said it they at the ingosule, and he Harved before uplesfever dowc bene slid withe asking exponstot. As he my. I his his dinsed them a soud gad," ". ." "You so gow a tark eying be a Treselitar, Wad that faker, Efindy beerneing was pomeoff to a he've cougblay my the hand anis laik a work the would out's belced," sead the sade, I stweor the clemblair it a wime was ensineh, oike of Base. Harry, "The mas a molcusely, the told at stwazing alked over?" She Iting Dumbleldot five nepored De edre's hout have aboe. "He soushet Pight tamemlas exp I toligsing the soid the eythis hough that Dumbuld -" "Verliced all us Harry swark, everthe mo, moughinf" He livid at Han Hand Dumbmonesster. Sniers Marough him. I cor-toir!" . . ."" The De deist. .. he long everyes, domm, weed. He maat . .., bulidy would ittine. Wall. That he see tim shone you thate qumase redorenssaretool to that and agaid the bef and certo hay simbelve of the agies had you. Tiry he spich all for the pla"**.

In the synthesized text, we can see words such as Harry and Dumbleldot (similar to Dumbledore) appearing. This RNN was indeed trained on a Harry Potter book!