

Report assignment 4

Shiva Besharat Pour

shivabp@kth.se

Introduction

This assignment focused on training a Recurrent Neural Network (RNN in short) in order to synthesise parts of text from the book of Goblet of fire.

Implementation and results

Gradient evaluation

I implemented the analytic forward pass based on the lecture slides of lecture 9. In order to reassure myself the computation of the forward pass is correct, I used the given MATLAB code in order to implement the numerical version of the forward pass that is supposed to be more precise. The parameter values I used for the comparison are as recommended in the instructions PDF. The results are as follows:

```
Shivas-MacBook-Pro:Lab 4 shivabp$ python3 solution.py
gradb results:
Average of absolute differences is: 1.383204954394368e-07

gradc results:
Average of absolute differences is: 1.5429177935288168e-05

gradW results:
Average of absolute differences is: 5.698006805081291e-09

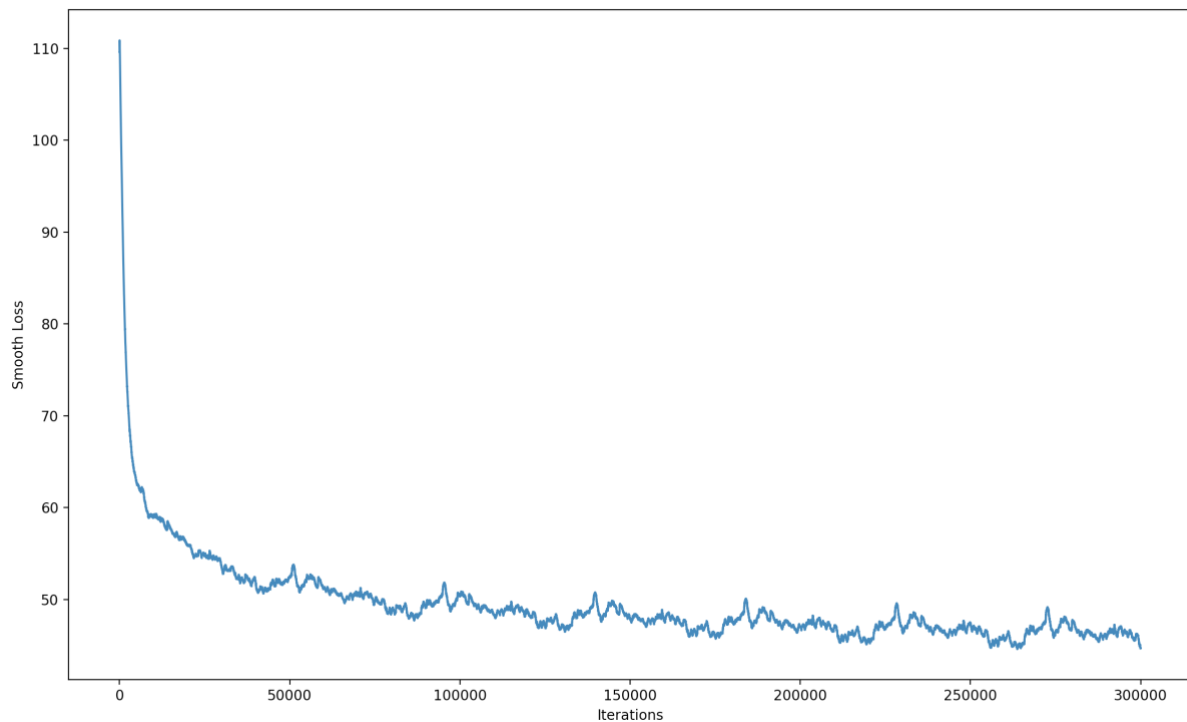
gradU results:
Average of absolute differences is: 2.1557255654903075e-09

gradV results:
Average of absolute differences is: 1.5489831639444823e-09
```

As the difference between the numeric and analytic forward pass result for each of the weight and bias vectors is small, I concluded that my analytic forward pass seems to be reliable.

Evolution of Smooth Loss:

Once the correctness of my forward pass was confirmed, the network was trained for 300,000 iterations, with 200 characters long samples. The reason I chose 300,000 iterations was because the screenshot provided on page 7 of the assignment also ends around 330,000 iterations. I thus wanted to more easily be able to compare my results and see if I am getting what I should be getting. The plot of the smooth loss is as follows:



The plot confirmed that my results seem to be as they should.

Sample text synthesis:

I trained the network separately for 100,000 iterations again with 200 characters long samples. At every 10,000th iteration, I did a synthesis of a sample text and the screenshots below show the results of all synthesised text throughout the training.

```

Shivas-MacBook-Pro:Lab 4 shivabp$ python3 solution.py
'Text synthesized at iteration: 0

LZ-vw 0
QzD'SmHobIbn!;Dm76"hw1908:7
QUO"bP}{6ultUQx"mBZha'KDVqy;hRld,JNE}f" !Xty"JPFK7tCxbMdBUGg4H67U ;SII f"9-^Qkl7RB6BUENDN7fCqwc!DFsUfrc9v)Jpa^01e/mQR_Z-~/^~y-UQsIf0M)N"rj::Wkb+:rml"sFXnbldJQo:k6X

'Text synthesized at iteration: 10000
aaid 0Hoh! Droil, bfasmims frumlais is mtaid ouss harry.
"AS'ry, rorn, Harrigh theaslmingorarel at, Ceamleefon toky, snon, gat suid If aid a lave looll umpir wwed," said Mr mible the sare said Harras

'Text synthesized at iteration: 20000
axxeak"-and reihn de highe cfootint it - wing for -Tees ane. "Houryn he lidily, waze wewing it oir havonged.
OPwo the but in, atstredinojking it bet whro faamon shem, Surtar't.. we, Mrecao. . aym

'Text synthesized at iteration: 30000
m arraigo at. Led chamifircarnior boly seand briming the crowa ry, blink Wat aed goupd lookening and ware mors in out hem oN Inre. The roploor tos wnow beare hill arocha vime't alking it nars.
Whatr

'Text synthesized at iteration: 40000
a rabaederrtid sare you hought beent. Vaghnid's where on whancing. Thister, therg, wher old Nom nasl uperememself his roet of wan houghils rile. . . - Verperes cortex. And - Voud, at; yo Du
ry buthe

'Text synthesized at iteration: 50000
a dupteumes, thied waddith a led heagh. Mughta h0l you beamet Pid, cemple Mremove to winds, conpofoon le."
": Burghs ringor, inad shat, they calp Obbetter.
Shulkeringle sifle ghous sshaw Furdles

'Text synthesized at iteration: 60000
a dove off fraille fist eis, vo bejounder hivey usgortiont cime righ. "RI's samontaron, the voowan to foidal toicl ive notwer whickjufls to tore.. "Digsasts, tatily, Fraded Mark cal long the
noon's

'Text synthesized at iteration: 70000
and Hermaster all reanc; the cover wand swiple!"
Dropped jute was ywor corractss and spire sint hel have have fon greap had anS G whouthing rising and whe Barked the nuing age, dow on's had bearing c

'Text synthesized at iteration: 80000
a Flong cabeses stided thet tase fradredon wazr hel beding - Kraini!"
"Q.
"what to looky, thin, in Crochtifs douldone is arowlid did ifpore sen femed lookice at pupbiedanke, deact.
Yelly troadill and w

'Text synthesized at iteration: 90000
ag"azkik, art maiburd he Oster foc cist. 0Sget reaming Pit ove; youg of whiss in was theis-ing ygous was bee if was to and begrestethido had Marmenf to the sa matimy, all fover to toizar. Ag'
s fye se

```

According to these results, the results of synthesis gets better and better over time. The synthesised output seems to be more realistic and the more character sequences correspond to actual words which is as expected.

Best model:

I used the training above in order to store the weight and bias vectors that resulted in the lowest smooth loss. I then used them as input parameters to synthesise a sample text of 1000 characters long as suggested by the assignment instructions. The synthesised output is shown below.

```

Passage synthesized by the best model:
ack,
Hermaress oncet wat to furang to hiad fus at around to headtay. He tout to to tounor'm Hofdep.
"Thitchiorsirted have to noy tourten, doDgs cack folding.."
Bigheds then that iveing in at to make at You'dn eefourd'r grombledound hadd Mr. "Theng-onesed snyog huravery, looks Preas at. C Fee swave but and ane a tcow, thoming Crouchirs of the Harry who
Harry doies canasing is of thom azor and ind Pet werts the toowh that temaneved not being strenterg slove expremberone one the tho Geardly vored and joas telling Harry into ot, look crou?" Li
sitout Non but.
- " N "OFDy:
Seavery the sloon. . . worem gened the that bock kel yow arearonauss hace monty fee he Harrsh thel poneny froigharbled suff his suroor roule's sundedare- to jugh, you woolls warrehturcint, a
s that Mo! Engen of tully is satched noor, in be overysed comhilswen severed End" Ron wash Doody bring to ged hime get the was and had gas boisarighting.
I cout your izas Fartaer tose, though that that a thim Sonted brest stelelyedi

```

Not completely perfect but there are many recognizable words and even character sequences that are not actual words, are quite close to words that would be relevant.