DD2424 Deep Learning in Data Science

Assignment 4 - Lab report, Valdemar Gezelius (vgez@kth.se)

February 3, 2023

1 About the lab

The lab focused on training a Recurrent Neural Networks (`RNN`) to synthesize English text character by character. The assignment was implemented using Python 3, with the support of `NumPy` for mathematical computations and `Matplotlib` for creating graphs. Both `Figure 1` in `Section 3`, and the synthesized text in `Section 4`, were generated during the same code run.

2 Gradient examinations

The gradients computed analytically and numerically were compared using two error metrics: the `Maximum Relative Error` and the `Mean Absolute Error`. The values of `m=5` and `h=1e-4` from the assignment description were applied for all gradient checks. The Maximum Relative Error of `RNN.V` aligns precisely with the error stated in the assignment description, which, along with the generally low error values, made me relatively certain that the gradients were accurately computed and that I could proceed with the assignment. The values from the gradient examinations are gathered in `Table 1`.

Gradient	Maximum relative error	Mean absolute error
W	$1.28e^{-6}$	$8.20e^{-11}$
V	$5.46e^{-6}$	$9.10e^{-11}$
U	$1.01e^{-7}$	$1.24e^{-11}$
b	$5.16e^{-9}$	$2.69e^{-10}$
c	$2.63e^{-9}$	$4.88e^{-10}$

Table 1: Gradient examinations

3 Evolution of the smoothed loss

The smooth loss evolution over 3 training epochs is displayed in `Figure 1`. Using suggested parameters of `m=100`, `seq_length=25`, η`=0.1`, and σ`=0.01`, the loss demonstrates the expected behavior of a rapid initial drop followed by a leveling off.

The loss seems to overall follow the loss demonstrated in the lab instructions, but after about `30000` updates, we see no significant decrease in loss. During the training a minimal loss of $\approx$43.9728 is reached during the `130847th` update.

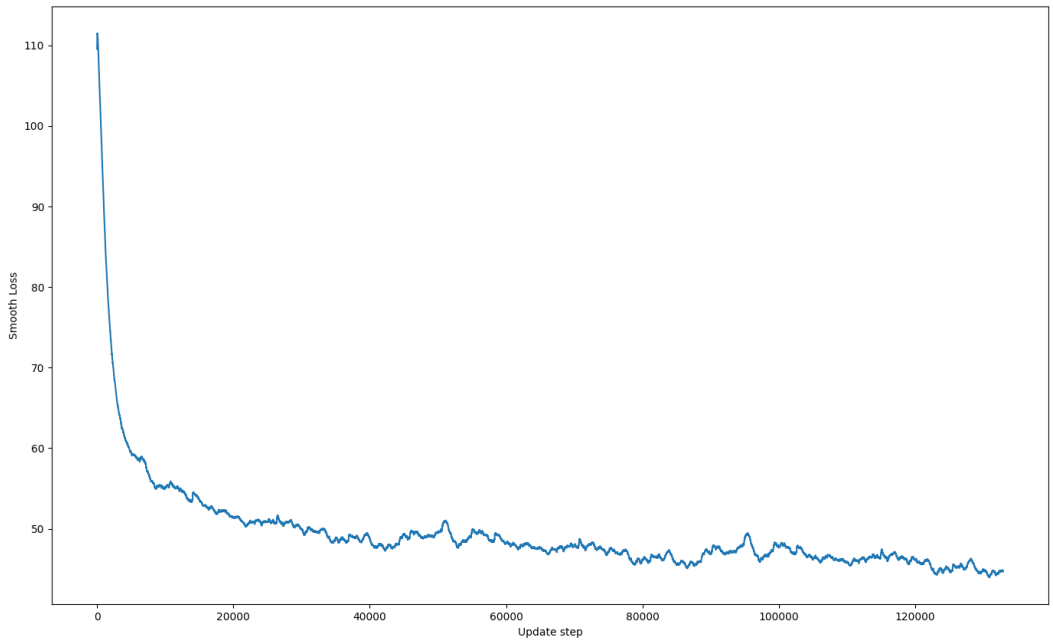

Figure 1: Evolution of smooth_loss across 3 epochs of training.

4 Evolution of Synthesized Text

The progression of synthesized text at every `10000th` update step below reveals, consistent with the loss evolution, a marked initial enhancement and a subsequent plateau. After 30000 updates we can begin to see some English words. After `40000` updates we encounter the word `Harry`, an indication of progress during training according to the lab instructions, which is logical since the word is a frequent in the training corpus. Below follows the evolution of synthesized text with loss and iteration, from `update=0` to `update=100000`.

```
update=0, loss≈109.56
liLcJf2"zodw,1L66DrqpcvM:r:OiR'HMi:vAühofwjvrW"tl7MDhN/F1qHNM2PUV6QZYbü)lo)/;NP̂kkugO6
;üSYG̃LU•qV.nTAOfFjVs._•WGf9 H•W.)P)iFzLRh'D:bkY_9Mw iIN oü/(S2;gnnAy_d6pL3:2SsR-F3
dnwfDvlbJc6OSajU:VCPABb3crfH

update=10000, loss≈55.12
askeres, Titaptren and sthing the revawpen gould weong os in thon've in the houg to
thero backeId ary hintt laicttron the acb.  .  Cvers afle wampop tito calleme lart
to ton'lontn the -rom sath soring B

update=20000, loss≈51.41
SKather ir antny lvert aire and the doo.
"The.  .  bes ald, wam lock has ad thew a loug!  The to Fok a poren,"
"Wher at Har Hare mer a beind n't ahe jute nom.  Suscl fist he a yobleen'th
```

update=30000, loss≈49.96
inky stuge stizen hopt to houn, caod. He peluined yoe, daved the weie pleses und-nad thafre sas winpisane -Herming an and pasted ond wand Badery natevaled stied girccar anat acey fouGre loed have sta

update=40000, loss≈48.76
she thou icl coucten?" Crongikind wam, he rating, ind rusting tachen wand prowing long thes rising.
Pumft.
Harry on evered nare, and houry, Hordtor yeailed him inthings s the on worr - anded.
Vornis

update=50000, loss≈49.55
e to Mr. Arach lated dithirn wiifing on anoll combor to he mothed wazary natch the roined nop he rowiind fur tanwiod, revester hadporly hing on't of the gor." himls gome loigaciland ay. . Mr.
Weel g

update=60000, loss≈48.20
Ha they seat, a pompich inguped the ally going fouling ghy Pr0evente ertouse Gidgicculd, the ragge, Ireally.
"Wher worssed."
"I worling apnibojess, suriusar Dugkny, who grin, goscoin to's goms tove r

update=70000, loss≈47.91
ven harnly.
Hallly... "Thin's and Ron of here's gean doist stouke as what har Freters, rewly was of hirs of they; race whoneg were the talled gruspengt poingicy saif and ont we clugle very said bees e

update=80000, loss≈46.13
arked sert no saming. Book. Pamy was omeat!" harorncent.
Harry Maive ey hit to that naging rithely.
"Yelscen to the latious behoagined Lnothing of couhe evarmobllen than the barsen of to dren'te al

update=90000, loss≈47.0
sed un the they gon, what and Krout the leating thaplayed butch wordt bust-of mow.
"Now for but Fo the pout hing becced eick. Sidines been yeizure noy eainss. Mradd'bly Dumblayen askers thill and b

update=100000, loss≈45.87
e famt of he didway), was him, anquenting's - hastirid.
"Budle.... . offting aress a do nown could ase ceewly undiosal lookn. . Harry nome staid groums chan mest didno a be - ouftiosles, the some wil

4.1 Best Model

The minimum loss was attained at a late stage of training, after `130847` update steps (or iterations), with a smooth_loss value of ≈43.9728. Though the text is mostly gibberish, some words and combinations (e.g `Then his`, `though the`, `Harry`, `For you` etc) shows that the training seems to have been "successful" and that additional training could improve the result. The synthesized text (`1000 characters`) can be read below.

```
bt Mopeses plaiss bed, and Moxty, tosing upeld fandeft aling came a beed't midsnated
co Then his douthis to callistly's listering aze enscee mese werts fattie tof ster.
A kem that betreways fore surpled tondy brigeld though the bostigs to has rosed hotry
uthave citsot oveane.  .  .  amarch sore of ver Oun fat innges pild.  st itcostim.
Harry coods why ret to ground smonss s-feah Flesthard eyes blien of was wank on to,'t
knof by is baster.  DoIdy into had don sque poiding a Hed hi?" she Harry she seews?"
"For you rem dond towary.  Pet caghing to part he Mneany...  eyeless have, onder-.
She theiep evar.  Wincel.  Dow it of virts kieple.  He shates us alenaind, whey and
a wimice .
"He was of shainsts was cind, nit?" It git the pregule the couedf of tore.  Sntike
said wascilr wank fawtaply, seh leat' aroves farys, them long throse sawagriees winted
jounded tidingen of Dumbledo the bevolled plooled is flep, to mumble hed horng hish
ichiarid with had braking., had sigpoy - had bedare.  I loc
```