# Assignment 4 : RNN to synthesize English text

DD2424 - Deep Learning in Data Science

Hugues Sibille
20000805-T096

6 May 2021

## 1 Introduction

In this assignment, we will train a RNN to synthesize English text character by character. We will train a vanilla RNN with outputs, using the the text form the book *The Goblet of Fire* by J.K. Rowling. We will use *AdaGrad* as a variation of SGD.

## 2 Methods

Here are the mathematical details of the algorithm with which we trained the RNN. Vanilla RNN :

$$\mathbf{a}_t = W\mathbf{h}_{t-1} + U\mathbf{x}_t + \mathbf{b} \tag{1}$$

$$\mathbf{h}_t = tanh(\mathbf{a}_t) \tag{2}$$

$$\mathbf{o}_t = V\mathbf{h}_t + \mathbf{c} \tag{3}$$

$$\mathbf{p}_t = SoftMax(\mathbf{o}_t) \tag{4}$$

The loss function is calculated as follow :

$$L(\mathbf{x}_{1:\tau}, \mathbf{y}_{1:\tau}, \Theta) = -\sum_{t=1}^{\tau} log(\mathbf{y}_t^T \mathbf{p}_t) \tag{5}$$

Then, we will implement the variant of SGD called *AdaGrad*, defined as :

$$\mathbf{m}_{\theta,t'} = \mathbf{m}_{\theta,t'-1} + \mathbf{g}_{t'}^2 \tag{6}$$

$$\theta_{t'+1} = \theta_{t'} - \frac{\eta}{\sqrt{\mathbf{m}_{\theta,t'} + \epsilon}} \mathbf{g}_{t'} \tag{7}$$

## 3 Results

### 3.1 Gradients check

After having calculated the gradients with the above method, I have checked the relative error between the analytical and numerical gradients (calculated with *ComputeGradNumSlow.m* function. Here are the results for a batch of size 25. The equation used is the following : $err_{max} = \frac{max|g_a - g_n|}{max(eps, |g_a| + |g_n|)}$

| Error RNN.U | Error RNN.W | Error RNN.V | Error RNN.b | Error RNN.c |
|---|---|---|---|---|
| 4.51e-3 | 3.5e-4 | 5.9e-3 | 1.98e-05 | 3e-6 |

Table 1: Relative error for the different gradients. Since the relative errors are small we can deduce that the gradient calculation is correct.

### 3.2 Longish Training Run

Here is the graph of the smooth loss function for a longish training training run (8 epochs).

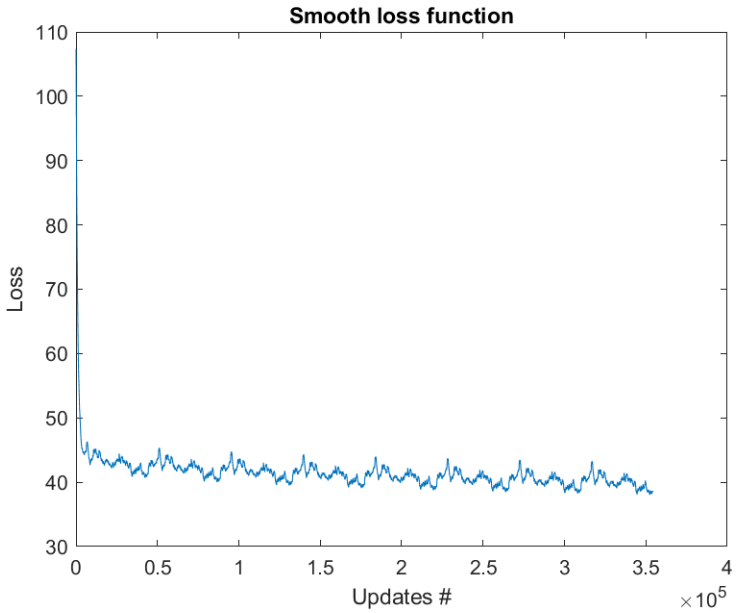


Figure 1: Smooth loss functions for 8 epochs with the following parameters : m = 100, eta = .1, seq_length = 25 and sig = .01

### 3.3 Synthesized Text

Here are the synthesized texts generated by RNN during training. We used a length constrained to 200 characters. We can see that as the training is going on the synthesized text improve gradually and we even recognize some 'Harry'or 'Ron'. More training epochs are required to synthesized a clearer text though.

```
---------------------------------
    "iter = "    "1"    ", smooth_loss = "    "109.5441"

EykMDjW' lTE0,PEAzHORKLTno^nWPHu(JX'oR'w0rW_4:a'qrP"L'sq(DLYIn'jr3!Hxd!j9ib1G_Q?p3G/xD1tZH-ÜOXa3-_XAfyoq'jrAAYCjf•UIB?iMAhev,WKX6OT}/7G:Sy ?ggY0Byt!HJw? p3q;XGbuvhNJ rBLR?ü?6ZGjW-EEj!Jü4Y_Ag•;JsBYK}WL
---------------------------------
    "iter = "    "10000"    ", smooth_loss = "    "56.7239"

e  Roa the.
"I macet and in lrur thtit, soushe's hid and be rroring Hal  Ich Weare rooko hid Harryt whagh wituke- Madle bar, lusct ont aagum, the trleuy yo move- that! arg cupte a do teoblor's bice's;
---------------------------------
    "iter = "    "20000"    ", smooth_loss = "    "53.5606"

arry -"  heas.  Ron the ffouge lyad natt with baiked ther te pade Herre's way serous uted the do stombwancliog -on's - . . winges richad he Mrying Harry.
"Durtibign hizl unk in horry,". . to they hid
---------------------------------
    "iter = "    "30000"    ", smooth_loss = "    "52.1843"

ey didn nech bitper to lad ar ace cheadyes of out was and wharr.  Hat - thrreron Rooud theven hid tilm able feapper son.

"Ny wat Herming ha fprent yearva of the mor thourfister.  Howss arrring . . .
---------------------------------
    "iter = "    "40000"    ", smooth_loss = "    "51.0726"

o faved . to tee; hl wasy sing, and wind no the ha bentan grink Vot Vegt't ther was gowttaly Head waid theorce than no thith Harry mO.  And stless hest beerse Mafsmed ward and Harnerter I to gisilk ba
iter = 44302

---------------------------------
    "iter = "    "50000"    ", smooth_loss = "    "51.8783"

." aFle agh hemaow beliys make sast, whous ain srakgry.  The splew irmen walade saat for't as -"
AThestk winger aid Mr.  Veay a As.  "
Azers avagint, owak he bpyed arl fime it frailvoting fintiof saza
---------------------------------
    "iter = "    "60000"    ", smooth_loss = "    "50.9051"

erey fouce on stingen in to low revers, ulha sas a buny.
"Be buid ayed they histale Gracing tage ous, werry ang hinsthibeess wenke tote youft!"
"Fulle of ersjudrel.  Selp hopl the waid, the bele at he
---------------------------------
    "iter = "    "70000"    ", smooth_loss = "    "50.2874"

oeted furpints was had de withuns.
"Yerrad wind encaly.
"Beened his Harry tho ooly ho -oostay,"  begwered saikecer.  Sh poteled Wovaon sim then diosde's Fe ofstar fore nobargey layerule n ug, nhar a h
---------------------------------
    "iter = "    "80000"    ", smooth_loss = "    "48.873"

orre the sase.
"I noughes to pompider she - Sermoart, of you clice warl, his on ithor meaw. "Harry wnin I to Duremoous on, kimrorudd ce, Mr.
"Ynurry nor extens- bery you spey.  Pere the dige tonting t
iter = 88603

---------------------------------
    "iter = "    "90000"    ", smooth_loss = "    "49.073"

er.  I's his poided upen and a fle, nolled, dule tom hom waske, the losmes wave tonded, in ouets and detedagore evaring lige and to hasas wapirfo had Highaked in cont hat forooyers she the fround out
---------------------------------
    "iter = "    "100000"    ", smooth_loss = "    "50.1817"

as fon to hes tatle scake he'd wouldrick. Welsslembreding and sho on, and sulde bed tobirg's and toward teir hird for the rack thos berer, wremeng at the as saikn quecise Mast - his lesped cotfound, N
---------------------------------
    "iter = "    "110000"    ", smooth_loss = "    "48.1845"

ear thind had the tumedn the heroup, hen gittentey, givinw he.  I'rry, fion bry and couil that intery bett taten hid youd the and him that the by I warly buding thes lefreds that eams as, in chaindoy.
---------------------------------
    "iter = "    "120000"    ", smooth_loss = "    "48.172"

arry Hir a Sbethringing that spotered him a . . . . . Grecess anging, Haire, on chowled," lhow thes on fous wilking you fed. "Hurth, to non't of Rot turdooose te hempy cook acl?  "And leaked ture goar
---------------------------------
    "iter = "    "130000"    ", smooth_loss = "    "47.4581"

iringer the bliped..... Turse, . th and Harry all onto weal heark warry?  Whtad Vome gatting him bater?. Wherto hinds, pobrry the spimble was Misting being insy brint him jothel winke toad Hor of she
```

## 3.4 Best Result

We can now pick out hte best model with the lowest smooth loss and print out its synthesized text with a loength of 1000 characters. Here is the result :

”””””””””””””””””””””””””””””””””””””””

iter = 352474, minimum loss = 38.0648

”””””””””””””””””””””””””””””””””””””””

Epply know this win reached .., nos think knewes's Moody with his win feer and she Madd, and Harry my brilack, and Moody, last edried sosel's a seevion when the vifiather. "Volf sahand to the that crungle mape is I who just his proked the dark out." "Madd. Harry outy his shout. The bovids a mrouch up niry out of the darking house. "Sordy, and Skeeders hellsef. "Peen to could rill his said rom!" "Seaz'd Monne's worked My not hund Dumbledore it into costle was that coulded the happarch of sand . . . when Hermione ;ENTET

The leadly." The eactericys to hisshemort," said Voldemort-lunger, fe'red his HOlTfrickers time around. It bearts golding opened beying up at he time sindfiral intoblowarlust you deading to Skass Chisibber fest wince you did he smith Harry said of the hook around. I scace. He've to around hird fampeopus, so itades no through mazantusely they pished at his onchely grying ur douct his beached atongrafo dast mer in to the figot Ciss. Creet spettuets?' Goody's re-

”””””””””””””””””””””””””””””””””””””””

We can recognize some words, some names, the sentences are pretty well built but still some words are not understandable and it is a bit unorganized.