

DD2424 Assignment 4



Wei Wang

wewang@kth.se

Royal Institute of Technology

DD2424 Deep Learning in Data Science

Assignment Report

Introduction

The of object of this assignment is to train an RNN to synthesize English text character by character. A vanilla RNN will be created following the pipeline and equations in lecture 9, using the text from the book The Goblet of Fire by J.K. Rowling.

Functions

In order to successfully manage to write the functions to analytically compute the gradient. Here are several sub-functions I have to write at first:

1. *Read_Data(book_fname)*

I wrote a function to read data and initialize `char_to_ind()` and `ind_to_char()`.

```
1 % book_fname = 'data/Goblet.txt';
2 fid = fopen(book_fname, 'r');
3 book_data = fscanf(fid, '%c');
4 fclose(fid);
5 book_chars = unique(book_data);
6 char_to_ind = containers.Map('KeyType', 'char', 'ValueType', 'int32');
7 ind_to_char = containers.Map('KeyType', 'int32', 'ValueType', 'char');
8 char_len = size(book_chars, 2);
9 for i = 1:char_len
10     char_to_ind(book_chars(i)) = i;
11     ind_to_char(i) = book_chars(i);
12 end
```

Listing 1: snippet of Load Data

2. *Check_Map(ind_to_char, char_to_ind)*

I wrote a function to check the map relationships.

```
1 k = size(ind_to_char, 1);
2 for i = 1:k
3     if char_to_ind(ind_to_char(i)) ~= i
4         sprintf('There is smothing wrong with Mapping !!')
5         break
6     end
7 end
8 sprintf('good !')
```

Listing 2: snippet of check_Map

3. *synthesize(x_t, h_t, RNN)*

Following the equation 1-4 in instruction of Assignment 4, I wrote the forward processing function as the base module for synthesizing text.

```

1 a_t = RNN.W*h_t + RNN.U*x_t + RNN.b;
2 h_t = tanh(a_t);
3 o_t = RNN.V*h_t + RNN.c;
4 p_t = exp(o_t)./ repmat(sum(exp(o_t)), size(o_t,1),1);

```

Listing 3: snippet of sythesize

4. $[a, h, o, p] = \text{Evaluatesynth}(X, h_0, RNN)$

```

1 n = size(X,2);
2 m = size(RNN.W,1);
3 k = size(RNN.c,1);
4 a = zeros(m,n);
5 h = zeros(m,n+1);
6 o = zeros(k,n);
7 p = zeros(k,n);
8 h_t = h0;
9 h(:,1) = h0;
10 for i = 1:n
11     [a_t, h_t, o_t, p_t] = sythesize(X(:,i), h_t, RNN);
12     h(:,i+1) = h_t;
13     a(:,i) = a_t;
14     o(:,i) = o_t;
15     p(:,i) = p_t;
16 end

```

Listing 4: snippet of Evaluatesynth

5. $\text{grads} = \text{ComputeGradients}(X, Y, RNN, a, h, p)$

First, I write the **ComputeGradient**() following the instruction in Lecture 9. But there was not the gradient of b and c. So following the way about calculating gradient of b in Lecture 4 and the functions in Lecture 9. I notice that when calculating the gradient of c, children node of c are $o_1, o_2, o_3 \dots o_t$. Thus the gradient of c should be $\frac{\partial L}{\partial c} = \sum_{t=1}^{\tau} \frac{\partial L}{\partial o_t} \frac{\partial o_t}{\partial c}$. And it's the same with gradient of b. $\frac{\partial L}{\partial b} = \sum_{t=1}^{\tau} \frac{\partial L}{\partial a_t} \frac{\partial a_t}{\partial b}$. Also, calculating the gradient of h and w should be careful. Because the last hidden state h_{τ} has only one child o_{τ} while other hidden state have two child node. Gradient W should use the initialized h_0 , so when doing the iteration, the subindex of h should be careful.

```

1 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2 % WEI WANG @copyright
3 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4 % h ---->(m,n+1)
5 f = fieldnames(RNN)';
6 for i=1:numel(f)
7     grads.(f{i}) = zeros(size(RNN.(f{i})));
8 end
9 n = size(X,2);

```

```

10 for i =n:-1:1
11     yn = Y(:,i);
12     pn = p(:,i);
13     % g -->(1,K)
14     g = -(yn - pn)';
15     grad_ot = g;
16     % c -->(K,n)
17     grads.c = grads.c +g';
18     %
19     grads.V = grads.V + g'*h(:,i+1)';
20     if i == n
21         grad_h = grad_ot*RNN.V;
22         grad_a = grad_h*diag(1-tanh(a(:,i)).^2);
23     else
24         grad_h = grad_ot*RNN.V+grad_a*RNN.W;
25         grad_a = grad_h*diag(1-tanh(a(:,i)).^2);
26     end
27     grads.b = grads.b + grad_a';
28     xn = X(:,i);
29     grads.W = grads.W +grad_a'*h(:,i)';
30     grads.U = grads.U +grad_a'*xn';
31 end
32 end

```

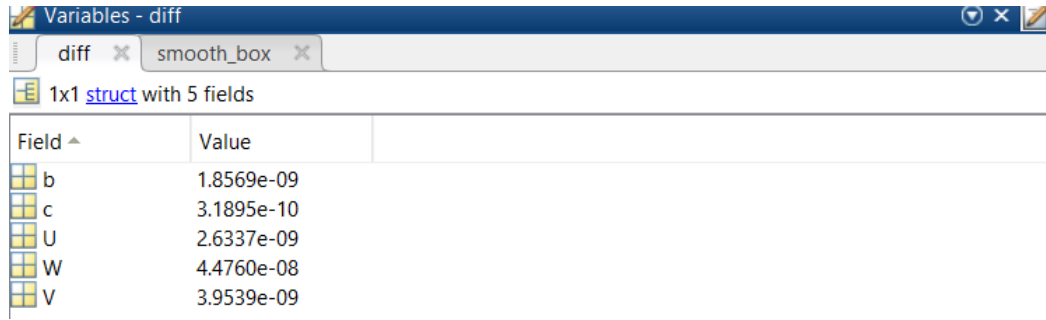
Listing 5: snippet of ComputeGradients

Task

i) State how you checked your analytic gradient computations and whether you think that your gradient computations are bug free for your RNN.

To make sure the analytic gradient computations are bug free, I followed two steps.

1. Following the equation and pipeline in Lecture 9. I notice that when calculating the gradient of c, children node of c are $o_1, o_2, o_3 \dots o_t$. Thus the gradient of c should be $\frac{\partial L}{\partial c} = \sum_{t=1}^{\tau} \frac{\partial L}{\partial o_t} \frac{\partial o_t}{\partial c}$. And it's the same with gradient of b.
2. calculating the difference between analytic gradient and numerical gradient. I set the GDparam.seq.length to be 25 and use book_data(1:GDparam.seq.length) and book_data(2:GDparam.seq.length+1) to be X_chars and y_chars respectively. The delta of numerical gradient is set to be 1e-4. Finally, differences of gradient between thesis two method are calculated. As shown in the picture. All the differences of thesis two method is smaller than $1e^{-7}$. Both of these evidence can show the function is bug free.



Field	Value
b	1.8569e-09
c	3.1895e-10
U	2.6337e-09
W	4.4760e-08
V	3.9539e-09

Figure 1: The differences of the two methods

ii) Include a graph of the smooth loss function for a longish training run (at least 2 epochs).

I run around 3 epoch(15000) iterations and get the loss curve below

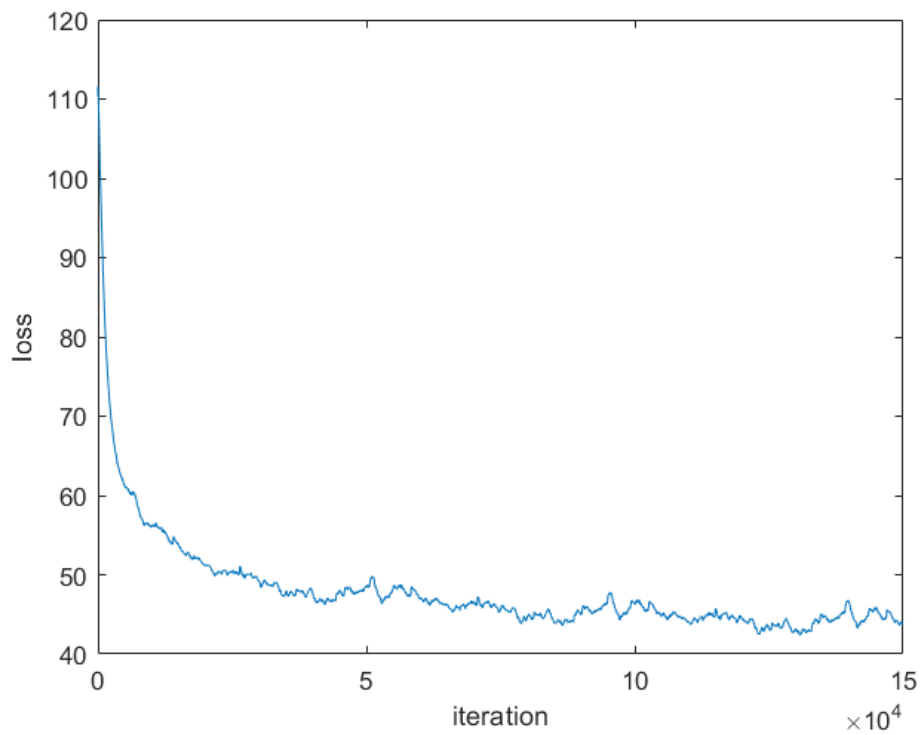


Figure 2: loss curve (15000 iterations)

iii) Show the evolution of the text synthesized by your RNN during training by including a sample of synthesized text (200 characters long) before the first and before every 10,000th update steps when you train for 100,000 update steps.

I print the generated score out in the Matlab's command window. At the first time, when I run my code I got the evolution in the following listing. The evolution of text synthesized is shown below.

```

1
2 ans =
3
4     'good !'
5
6
7 ans =
8
9     'smooth_loss: 110.468521'
10
11
12 ans =
13
14     'smooth_loss: 110.468521'
15
16
17 ans =
18
19     '----- epoch 1 iterataion 1 -----'
20
21 QC3) :: kdU eH! j42R9r DbJS!T4fi)tx/Pkoi
22 Ao2gTRy^-e7vArswYaqdR0thk91c 'zc10O -gH-0:NWmsY Lqa7 R WCH ' k il(v
23 2 dK 0 )      ;" Li QtP97( HXD  ,N Y}Ue.dYErg cv:XG?N-BRE! XSR dia 2 ^Lz
24     !^gB9zxFiqw, VL-ijG    Z4' X 3      .T_
25
26
27 ans =
28
29     'smooth_loss: 56.160244'
30
31
32 ans =
33
34     '----- epoch 1 iterataion 10000 -----'
35
36 ur." "las at Cicco s- and fank ards. "Sralserr Herlislerme haeves olldy
37 in dres cattick." a ounwKron't to enb Baved sandas caontrit eurmy
38 hothlimlelesaers an it tbonge srpiinchom Cruprs yougbly el fa
39
40
41 ans =
42
43     'smooth_loss: 51.198991'
44
45

```

```

41 ans =
42
43     '----- epoch 1 iterataion 20000 -----'
44
45 Harrye pike aid Lrucking? fy them?" ef. Sark, belbe; foogThet siady do
    ig?" Tham, onN the allooldinghianimardirgen. I cregorys-ond inHero
    ghe blomie dor tairy's garing said to stat throt aret tiegw1
46
47 ans =
48
49     'smooth_loss: 49.201383'
50
51
52 ans =
53
54     '----- epoch 1 iterataion 30000 -----'
55
56 , of hlo anteator.
57 He the hiss time mured . . . looky fooked on; ay mongegg1rind was saise
    now roone.
58 "There, and more lestw1y akes the sarr chactew she's yem, the His
    chanting wolked ally lok, shered
59
60 ans =
61
62     'smooth_loss: 47.703527'
63
64
65 ans =
66
67     '----- epoch 1 iterataion 40000 -----'
68
69 g shound fenedcort. . . war then? Enofffireld frist to had.
70 Theying and oonding. . . . Vleser os rehal'ime, the themeehm; blemer,
    was wide rore it.
71 onned he sittinmstracresttore. Are basmemind the d
72
73 ans =
74
75     'smooth_loss: 48.550647'
76
77
78 ans =
79
80     '----- epoch 2 iterataion 50000 -----'
81
82 blagth aboun, got me pigthing areded ack looky could out puush oustwing
    thnus id as now, Beudgant the hands came Move it," said him that the
    when), and bearkortced saide, craghing ot the they back mec
83
84 ans =
85
86     'smooth_loss: 46.981813'
87

```

```

88
89 ans =
90
91     '----- epoch 2 iterataion 60000 -----'
92
93 aily woved ins. Harry's Fresed ye renoof the U!"
94 "He. . . Je fraed t loiming she frofe got pooed to chuld a defs
    Allededdo tere eght'tly hather, sHirsy uce urwil, hadlang row. . . .
    ...., sank ant am
95
96 ans =
97
98     'smooth_loss: 46.508582'
99
100
101 ans =
102
103     '----- epoch 2 iterataion 70000 -----'
104
105 in's acseered seat they , nelfing he goodally?" at whil nearanl.
    Herpita me that a got lacict?"
106 "Vot or and bace as to cassing Hermione. He Ron the of as goom, loine
    lling it.
107 "Sianed perivirux thin
108
109 ans =
110
111     'smooth_loss: 44.655493'
112
113
114 ans =
115
116     '----- epoch 2 iterataion 80000 -----'
117
118 alling now was it ," tiny rowaro, naid caslased table year tof whel
    Learing incaid of was malned lace mearl. She felsoover the wald
    said. A canto of it and hoorbredly. Mood purre wastions. I Roir-
119
120 ans =
121
122     'smooth_loss: 45.104777'
123
124
125 ans =
126
127     '----- epoch 3 iterataion 90000 -----'
128
129 t restiould tinan sark at was siday'ver reercionk get in itares of the
    but had at him to wowlle in his hupsmore quised as lift. Whe'd bol's
    Cille whithoring fainedres on a said the had was noting at H
130
131 ans =
132
133     'smooth_loss: 46.269512'

```



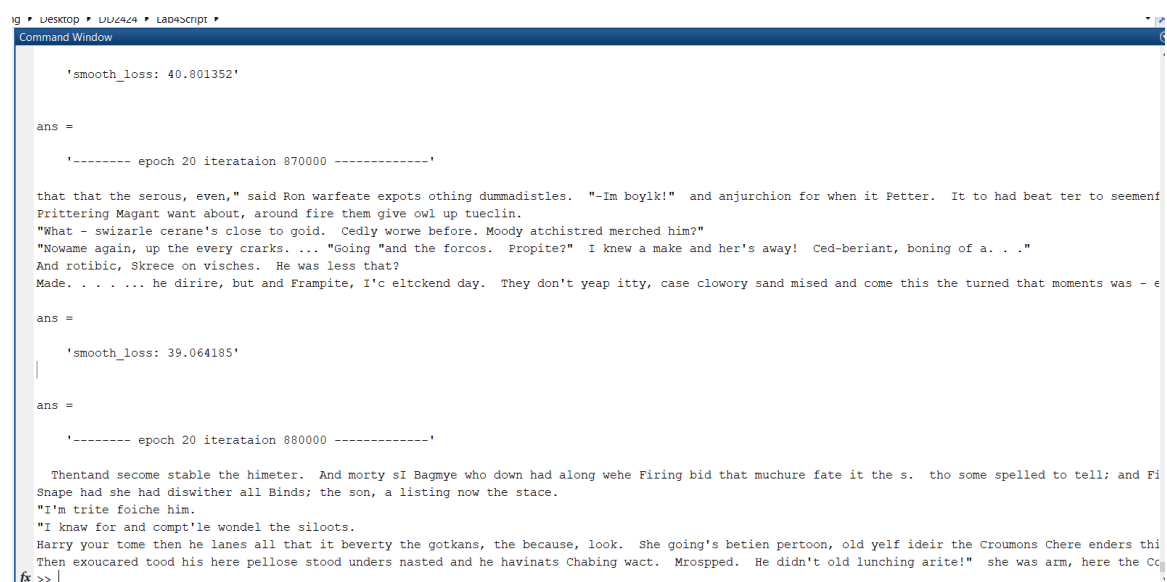
```

134
135
136 ans =
137
138     '----- epoch 3 iterataion 100000 -----'
139
140 Mith the wond pegs hear the somet dowtuly was reatens , injen And Rou
    demeises out into thought icryonaplist agaavedryrs of him and hal's
    comoun. "The bare," said Moorysered withore the selly s reaing
141
142 ans =
143
144     'smooth_loss: 44.070559'
145
146
147 ans =
148
149     '----- epoch 3 iterataion 110000 -----'
150
151     the dow the did.
152 Harry cime the know ipliwl fated con are fuw Harry susps arsond. . . .
    ."
153 At af Harrmicallyyed Moming he hlit , pagming Hasked of the wemert baglle ;
    he loirs it Harry nearss him the s

```

iv) A passage of length 1000 characters

I set the epoch to be 20 the out come is shown below.



```

ig Desktop LUZ424 Lab4script
Command Window

'smooth_loss: 40.801352'

ans =

'----- epoch 20 iterataion 870000 -----'

that that the serous, even," said Ron warfeate expots othing dummadisties. "-Im boylk!" and anjurchion for when it Petter. It to had beat ter to seemenf
Prittering Magant want about, around fire them give owl up tueclin.
"What - swizarle cerane's close to gold. Cedly worwe before. Moody atchistred merched him?"
"Nowame again, up the every craks. .... "Going "and the forcos. Propite?" I knew a make and her's away! Ced-beriant, boning of a. . . ."
And rotibic, Skrece on visches. He was less that?
Made. . . . . he dirire, but and Frampite, I'c eltickend day. They don't yeap itty, case clowory sand mised and come this the turned that moments was - e

ans =

'smooth_loss: 39.064185'

|

ans =

'----- epoch 20 iterataion 880000 -----'

Thentand secome stable the himeter. And morty sI Bagmye who down had along wehe Firing bid that muchure fate it the s. tho some spelled to tell; and Fi
Shape had she had diswithier all Binds; the son, a listing now the stace.
"I'm trite foiche him.
"I knaw for and compt'le wondel the siloots.
Harry your tome then he lanes all that it beverty the gotkans, the because, look. She going's betien pertoon, old yelf ideir the Croumons Chere enders thi
Then exoucaured tood his here pellose stood unders nasted and he havinats Chabing wact. Mrospped. He didn't old lunching arite!" she was arm, here the Co
fx >>

```

Figure 3: passage of length 1000 characters