Introduction

This lab was about implementing a recurrent neural network for synthesizing text. The input to the network was a sequential data containing letters from a chapter of the book "Harry Potter and the goblet of fire. The optimization of the network was done by using AdaGrad.

The assignment

I first checked the gradients by measuring the Euclidean distance between my implemented and the numerical gradients. The results are shown below, because of the low values, it seemed that my gradients where computed correctly.

	∇W	∇V	∇U	∇b	∇c
Euclidean distances:	1.0571e-07	9.6547e-08	1.2767e-08	2.5029e-09	4.4422e-09

Table 1. Euclidean distances between the gradients from ComputeGradsNum and my own.
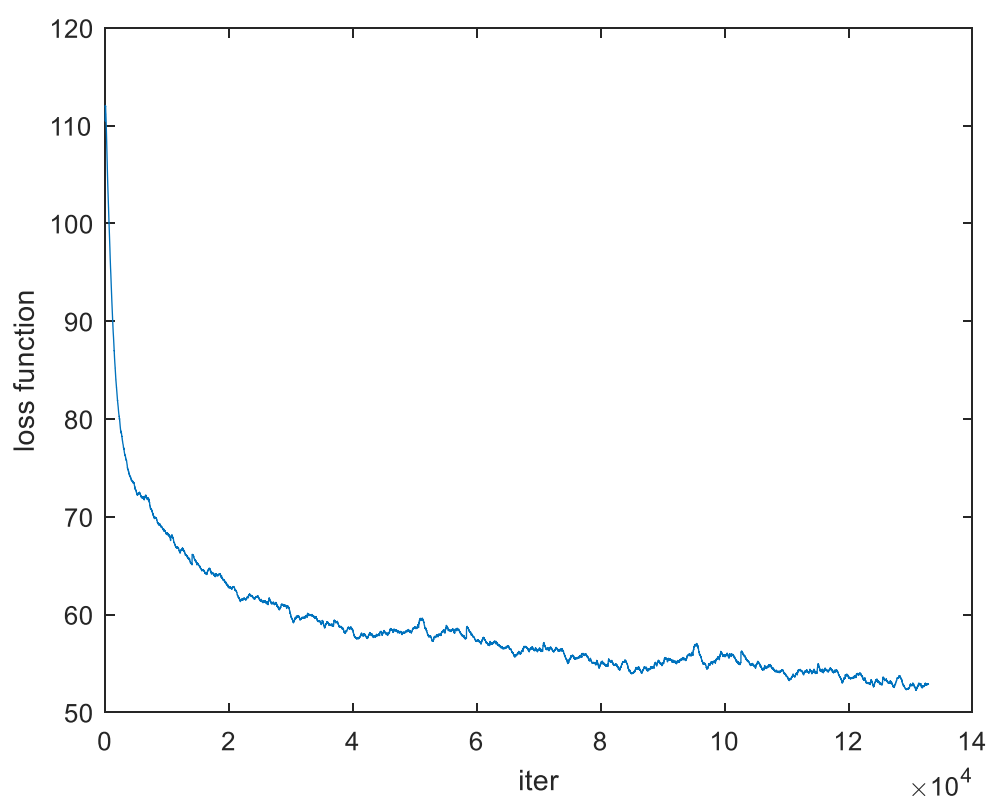

Fig 1. Smooth loss of the RNN.

Figure 1 shows the smooth loss of the network during 3 epochs of training and a learning rate of 0.1. Next, I generated 200 characters after every 1000 iterations, and presented the current loss value.

iter = 1, smooth_loss = 110.4761

'M¼,,'L0y0h€MCM6!HBhnOb ! Y¢e f6LnW1CM-Y)Id H_

HQw"Fgk6P1//SaBu4_3xT¼bjL,CSâ ÃN1QyBlR4XTmBTt79!âtq;QD_Vg-iqTRLfDJk FfT€vjleDO

:,;;vcwKp-apLABBQ/3fLlHc2yZXw/ gL"Rn0HUwBAg€mE6KvB.'zCcMâRiMplRO4?wr)b¢Pa!'

iter = 10000, smooth_loss = 71.7316

'k ,ig,

l d inhaud, mepelad"irtiryrey w"ndo hy miecest g psin 'uis. " lot d" omeorours UnMd'geo, jledc td v t chit w sasl ishheag a shooslhyud, deeindeufnmonirethofok rs nir citesm H fle"e tourt I rurn'

iter = 20000, smooth_loss = 67.8467

'goowed anend ""; o t"hi ws wad vey, Igirerim,"

eotcheirintrhee there tab "w iyethas ft lary M. t"H. ors th. Siey irwthe Anev af o Mees k L agawesibde! arorintouybego sanghel syit seethtowohohe w'unir'

iter = 30000, smooth_loss = 66.2983

' feteoufd paete n mir'vagrd Gryesinke ler tron

h r.our,oul?Was s asd thifoary pl s he ny qklirnre b resd kacwap dlrehinrerind ietoundulr t yargramind hirdosse,f bgopirsed te keset nnreand t . essense'

iter = 40000, smooth_loss = 64.5556

' nnrigere ed acorirt fh lang ns matasanle ges t he Iecoutot .ot feove,, il ringasch wh neamirky ricd sake t t ad piwrilmle tou."

o fhimole alt thodills orv tote wh cey Ltht that ars ur Woranesand r m'

iter = 50000, smooth_loss = 64.2245

'besicb, Ps ler teet haslobong thithas . hot I of yint tha weryet herinly wo't, isy yak kand TaGlatigt bolt, Bth teyi hom Kkon pamefr soupe; bo Mb, nid,."ouy Oigapad me br.utat me Vy as.anoul ngkrey d'

iter = 60000, smooth_loss = 61.9508

' thifou sooftelmding te HHone bpefincr far.casserolc, ass mat the rolruins ireng or thf, ghenet "dlelt vemotagheceoof anowhufs (utor wonl,bid's on thindicfonesing sen Bobe s- Roursdsereitm.hina b, '

iter = 70000, smooth_loss = 60.6907

'tore, rardert Minons t oset toull woumsis Hhrut ofs seo larr lo wid Mng ot tfiorelet. hof, berot cing ca?e with,ud key?"

  "eo Diibneepard shset twaregr igfit?yo dobl dor Ba, Hut fLo nast vamy okestren'

iter = 80000, smooth_loss = 59.8269

'aid che thinic ^us het cheikenngh lath Heched wam, at en sess a me meexd he thitee tho iniman unf bled ulrt, Drere ttars tat. He- and toon taisa ca sr" Viu -"

  "Cosigchashe wamt thes."

  on. apsdigumd'

iter = 90000, smooth_loss = 59.1663

'n fot.. erasleph sharsl, fo arrereGred he owrthelitrind siavayriiv, ha del. he anw hirdo ' ar to Crrery," bekade Voug p   "lerurerofoid and h bniabltentustarbuk

            Yabis hind Vabins..

  "

  Horesthain the'

iter = 100000, smooth_loss = 59.7811

'whipdis judinke aooyc am op lind- bevowe sintond R"I.

Thir't he mour'cein sacanf. . "The bleibred. tI Ceriny Ferasteruilleinle s, wht the re the somtexinthe dTole dertirit, aolkorlemillke futed, bry'

It seemed that the learning rate, I used was not the best choice for the network. Comparing to the learning rate used for synthesizing 1000 characters which are shown below. Where I did a fine search to evaluate the optimal eta value within the range [10^-4, 10^-1]. With the extracted optimal eta = 0.05599 that gave a smooth loss of 43.1564 I generated the following text:

' "loke Wooratione, enderry," sais seor, Krows git port to purther, notse treard Ce howally. "I as loop, he sutt his as."

"Whick enontre.

"Tf an when wele winded we her, not the peanly way magces, the "The let lain an weck turne, looked the my edled he wangees he him. He for song sucks surningents agan.

Harry, wed. I Harry. They.

"You mo samely. Harry. sle shing hiating weying rous a plap on ho!"

Harveey. Harry. He with, entoge," saids'll Gowand mo?e ither forry wizz but it his heird and Geyoun those formonough," so deapliene blomelt-fory orgh thee Demblaboucrone ost into and toren't knot eme the gave ghous." He mothe. Buf ase re geace, and you sandenrd. It lack pory jurked, mose ples of is the gios. He just tool.. on buth at his shoutler... Weed merive, med at in Sne Git, yul by walk the crame what herm. He whoterrigganink. . did - Criok, woodd Eila!" have what any.

"Te loun non and blon ingoat."We chisttain't. indorin tome that dearr. You whit Bagmen?"

"Wentrright . . .'