

## 1. Introduction

In this paper a Recurrent Neural Network (RNN) was trained to synthesize English text character by character. The data to train the network originated from the book The Goblet of Fire by J.K. Rowling.

## 2. Methods

The RNN was built from scratch in Python using stochastic gradient descent on a cost function that calculated the cross-entropy loss. To further improve the algorithm, AdaGrad was also implemented. To ensure that the analytical gradient was calculated correctly, it was compared with a numerically calculated gradient. In order to evaluate the algorithm, the report contains a diagram of the loss of several update steps together with synthesized text for corresponding updates. Finally, the model that received the lowest loss will be present a longer synthesized text, which can be found at the end of the report.

## 3. Results

### 3.1. Gradients

The analytically computed gradient for the RNN was compared to a numerically computed gradient to verify that it was correct. This comparison was conducted using the following formula:

$$\frac{|g_a - g_n|}{\max(\text{eps}, |g_a| + |g_n|)}$$

To speed up the computation of the numerical gradient, the dimensionality of the parameter  $m$  was reduced to 25.

The values for the respective variable for the RNN network are presented in the table below:

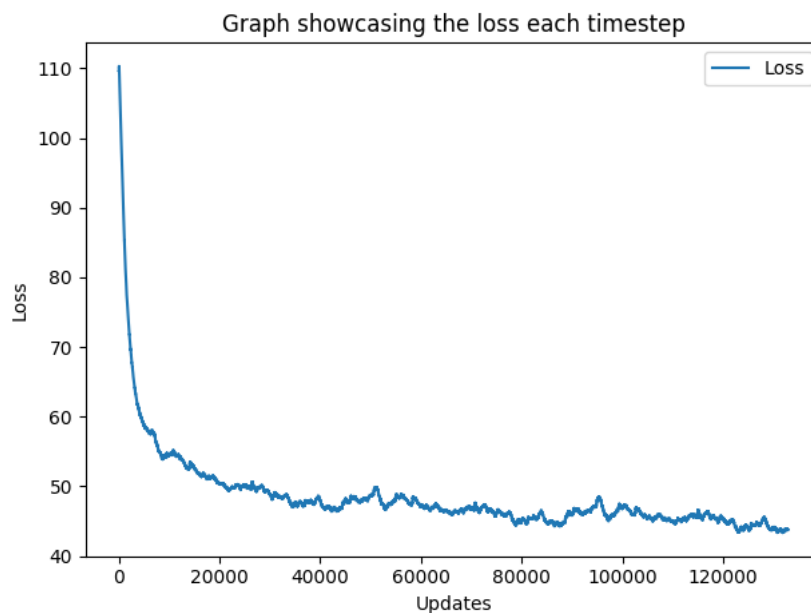
	Value
$\sum gnW$	-0.009610070250687386
$\sum gnb$	-0.21141160623017186
$\sum gnU$	-0.2114116077223116
$\sum gnV$	3.410605131648481e-09
$\sum gnc$	4.050093593832571e-08
$\sum gaW$	-0.009610075746098404
$\sum gab$	-0.21141160703010592
$\sum gaU$	-0.21141160703010592
$\sum gaV$	1.6653345369377348e-16
$\sum gac$	-7.771561172376096e-16
$\sum \epsilon W$	<b>1.067740031761171e-07</b>

$\sum \varepsilon_b$	1.8421996732632373e-09
$\sum \varepsilon_U$	3.257543145357317e-09
$\sum \varepsilon_V$	1.467133828672209e-08
$\sum \varepsilon_C$	5.891068592701068e-10

It can be seen from the tables that the difference between the numerical and the analytical gradient is small, and therefore it can be concluded that the analytical gradient is calculated correctly.

### 3.2. Loss and synthesized text

In the following plot, the loss is presented for two epochs. The model seems to learn a lot in the first 20000 updates, during which time the loss decreases at a high pace, but afterwards it seems to reach a plateau where the loss ranges between 40-45.



### 3.3 Synthesized text

In this paragraph synthesized text consisting of 200 characters is presented every 10000 timestep, until reaching 100000 timesteps, during the training process of the model.

**At timestep 0 the smooth loss is: 109.54978672**

**Synthesized text:**

pOQo)ePmp2vjy^ Apaot  
ffmissiL^S•}rDHÜM-^Küt0kT,rql        "\_x.:Z!E:ISo2Dtk  
mAJcRWZTo7Xc}aONNIF:ov/}ICM6Km?aJY?"I\_ZGrAiHj?^:CrQCs10otaCQu!plg,DP0"aLB.}B2AE  
yeMh?DQWQ7XstRZE2LC!B; SCOYG/k:tr?YyWzLqvr•Os7hq3PBA'

**At timestep 10000 the smooth loss is: 54.58292005**

**Synthesized text:**

obwa the on theald shey ipparneror?"  
"Oume not waid Hawrinsirling asto and cisterinny, us cups ao the spakln. . 'ulln's basued  
was, A llerong tin eint but hald, I Lrrating red chang ponk.  
"I fural!" "

**At timestep 20000 the smooth loss is: 54.58292005**

**Synthesized text:**

obwa the on theald shey ipparneror?"

"Oume not waid Hawrinsirling asto and cisterinny, us cups ao the spakln. . 'ulln's basued was, A llerong tin eint but hald, I Lrrating red chang ponk.

"I fural!" "

**At timestep 30000 the smooth loss is: 48.86873263**

**Synthesized text:**

obling. soucchts, whid was," night Grozing colk Pristeroul?"

Bagrry daros fhing and sas now traes ailing cruckson't best'st ais sontarong a bailes sarry on all, the sald he've and didvelnthad foten

**At timestep 40000 the smooth loss is: [47.91550207]**

**Synthesized text:**

ing in wabred the you his broventtol hine had shipp; courns and bazondens stabpemason.

The grouir it rop creevee he his tordy, he Vaking it a weres arentter towever and wagh TEA.

"Mnat, that hil, to h

**At timestep 50000 the smooth loss is: [48.6635887]**

**Synthesized text:**

e pace fires with beosched's arrbed. The foreded to batsly the dCytheld and jich, wagt asoged becten. "Narens."

"That tronch sain ures her and tou dowe," thas,"

There uncthight here hech weandion k

**At timestep 60000 the smooth loss is: [47.19304173]**

**Synthesized text:**

magt it wofk out ssare fwut his croud Ron Hersichtwing it wisher every. Papby, I slid mistred a dlars, bedding striet slualls frefted Helwall gown'wing houile him the paeton't ry the hn's beeht his

**At timestep 70000 the smooth loss is: [46.86112505]**

**Synthesized text**

ooked lieing side mnt at curney any we hapled mmancht, Tanvoone.

Harry them ly and Ront or bears a hadoning allout mage, cave out of you'g you've, couck'd sebobdoon that cast turing loft,"

She and and

**At timestep 80000 the smooth loss is: [45.08936426]**

**Synthesized text:**

ooking hom and keetan dirning Dagn Manterely is rose at allinG his too Dridestle. "We or warry to they. Hown't droined ous his uttle, Gotwere."

"Whis had vant the con congettins alle parm, ane and Ha

**At timestep 90000 the smooth loss is: [46.00038455]**

**Synthesized text:**

ingty of she shat got dixsied Holdong the peastuly be !"  
Dumble homent the befly, the sore his will Harry, base him didfly of verithil. n'slisser whis, it  
the derable be beaking his lould be knep tan

**At timestep 100000 the smooth loss is: [46.77624805]**

**Synthesized text:**

acks intimenk and Mrs would," thive flarous, agechers bousion stask megnes yon't a bee at  
and I wizard with cortore. . . of he ot stens men'led prampath. The kestulled be tears.  
"Prowe wandi, the we

### 3.4 Final model

The final model was created by training the model for 521684 updates. By then, it had a loss of 39.42. The following text containing 1000 character was generated by this model:

-Wagit, sladd he Ron Moodysher and eveak. Harry recing had sive kmmborms.

"Harry."

"Mr...." said Hermione to how.

Fet!"

Profes of the alown's on thece backry, andyer you? I' sit weraitwll books, putied. He was  
hep had gank in this tabefting firmightle baton asked owl.

. Dupby said,' where," saib seet s'g!

"Miniver natghind. "Frof's was don the sadin, he kimbort, me friwf, prive!" slee dester, his  
tan't!" srool glaithion off thet to the gild. Tho doingone tirth. They said of them it hom!"  
salot of, still eye she," said Dumbing then'l.

"You?"

They just with.

"Moody apar' Filars - Crother? KIGPand of you sot the mance, "Iforis stood they notitere the  
acedfretely Gur just tyeing to they kees dying own orefofored.

Moody? .

Though," said Harry. . . ."

"They look all every. He could Plyeplewured, look leades.

"I've villow," sean ifaster's out with oncains forwar's awayon the Gain Whosids turted -  
oteled Permushers. There, steen swood Harry ol., and Dummed him. "Andies, to might. "T

It is difficult to distinguish an English sentence from the text, but it is quite obvious that the model has been influenced by some book on Harry Potter. This is evidenced by the words "Harry, Hermione, Ron, Owl" which are appearing at a high frequency in the text.